Order Number 8918589

The linear algebra of interpolation with finite applications giving computational methods for multivariate polynomials

Olmsted, Coert D., Ph.D.

University of Alaska Fairbanks, 1988

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

THE LINEAR ALGEBRA OF INTERPOLATION WITH FINITE APPLICATIONS GIVING COMPUTATIONAL METHODS FOR MULTIVARIATE POLYNOMIALS

A
THESIS

Presented to the Faculty of the University of Alaska
in Partial Fulfillment of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

by
Coert Olmsted, B.A.

Fairbanks, Alaska
May 1988

THE LINEAR ALGEBRA OF INTERPOLATION WITH FINITE APPLICATIONS GIVING COMPUTATIONAL METHODS FOR MULTIVARIATE POLYNOMIALS

by
Coert Olmsted

RECOMMENDED:

J. P. Lambert

C. A. Lando

J. V. Olson

R. J. Piacenca

G. A. Gislason, Advisory Committee Chair

C. A. Lando, Head of the
Department of Mathematical Sciences

APPROVED:

A. D. Shinkwin, Dean of the College of Liberal Arts

L. Proenza, Dean of the Graduate School

4/29/88
Date

Abstract

Linear representation and the duality of the biorthonormality relationship express the linear algebra of interpolation by way of the evaluation mapping. In the finite case the standard bases relate the maps to Gramian matrices. Five equivalent conditions on these objects are found which characterize the solution of the interpolation problem.

This algebra succinctly describes the solution space of ordinary linear initial value problems. Multivariate polynomial spaces and multidimensional node sets are described by multi-index sets. Geometric considerations of normalization and dimensionality lead to cardinal bases for Lagrange interpolation on regular node sets. More general Hermite functional sets can also be solved by generalized Newton methods using geometry and multi-indices.

Extended to countably infinite dimensional spaces, the method calls upon theorems of modern analysis.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

Figures

Tables

Preface

In a lay level lecture at the University of Alaska, Fairbanks in 1983, Marvin Minsky said that he enjoyed working in Artificial Intelligence in part because "you don't have to spend all your time in the library to find out if your research is new." He pointed out that in classical analysis, one's striking result may have been written years ago by Hermite or Laguerre.

Interpolation theory goes both ways. New results are continually percolating out of analysis and applications while, at the same time, well preserved and useful chestnuts are being rediscovered every generation or so, sometimes clothed in bright new guises. Hermite wrote many and Laguerre may well have, but I can not tell you any of the latter, as my library time is finite. The bound, however, is large enough to give me confidence that some of my research is new and that the rest of it can be unifying.

It began with a practical need to compute polynomial interpolants in the plane for contouring, optimization, and finite element analysis. A geometric method for labeling a trianguler array of nodes was extended to general dimension and then general degree. Extrapolating from cases of lower dimension and degree led to the formulas of Appendix I. From these forms the general case was worked out and published in *Mathematics of Computation* in 1986.

This thesis covers the highlights of the secondary research which was necessary to place that result in the context of contemporary numerical analysis. Encountered along the way, the case of infinite interpolation provides an interesting library tour of modern analysis. Perhaps in the future such toil can be performed, courtesy of Minsky et al, by computers.

I use the first person singular and plural throughout. The former signals my personal opinions. The latter is meant to include my graduate committee and any other interested readers.

I would like to acknowledge invaluable assistance and crucial moral support from my advisor, Prof. Gary Gislason, and also Prof. Pat Andresen who led me through the ambiguities of linear algebra. Significant support was provided by my employer, the Geophysical Institute of the University of Alaska. Many thanks to its directors, staff and the several cohorts of graduate students with which I merged.

CHAPTER I

Introduction, Definitions and Background

1. Analysis

As approximation is a paradigm for analysis, so interpolation is a paradigm for approximation. Interpolation, however, is more primitive conceptually and historically. It seems to be what untrained numerical analysts intuitively think of. There is something natural about extending our knowledge from a limited extent into the full domain by smoothly filling in the gaps. Thus in some instances, interpolation may dominate the nomenclature (e.g., Forrest [1972]) but may be the most trivial of the approximation constraints. Beziér curves, a Bernstein polynomial approximation technique rediscovered by Detroit auto designers in the computer age, interpolate only in that their endpoints are fixed. More important in determining the shape of these curves are their practical mechanical properties and the useful parameterization whereby the curves can be controlled by a sketch(polygon).

Also there are computational constraints. After all the analysis is over, all we ever compute are polynomials. Some of these constraints are linear but others may not be. We will see that interpolation, with few exceptions, is a linear process. Also we want our constraints to be equality constraints, at least in some places. Thus for some (generally linear) method of comparison, the approximation must be exact.

An essential part of approximation is, of course, specifying the exact or "true" value—what is to be approximated. This is usually provided by the "customer"—an application in science or even pure mathematics. Since this is a thesis in applied mathematics, the objects of our study will be functional relationships between variables taking on values in a field. Usually the field is the real numbers R or the complex plane C. We will denote our field generically as K but often assume that it has an absolute value or modulus $|\cdot|: K \to R^+$ and is of character zero. There may be a theory of interpolation over finite fields but the library constraint prohibits its appearance here.

Vector valued functions are studied componentwise so that we may restrict our attention to

$$K^X := \{f: X \to K\}$$

where X is a vector space over K. Typically $X = K^k$. It is the vector structure of X,

inherited by the function space K^X which establishes the linear setting for interpolation. Within this setting the application "problem" specifies a family $\mathcal{F}$ of solutions. To maintain the important feature of linearity, we require that $\mathcal{F}$ be a subspace of K^X.

The next task of approximation is to render the solution computable by selecting a further subspace $\mathcal{H}$ of $\mathcal{F}$, the elements of which we know how to compute. Strictly speaking these are (as noted above) polynomials of specified degree in K^k. We will, however, in the spirit of analysis, extend the ideas to countably infinite bases of trancendental functions, thus enlarging the concept of computable to include primitive elements of approximation. However big our space, we must know that $\mathcal{H}$ is strictly smaller than $\mathcal{F}$ and more tractable computationally, or else our approximation is trivial and useless.

Having settled on an approximating subspace, we now need some condition which, given an element of $\mathcal{F}$, specifies an approximating element of $\mathcal{H}$. Thus, at its most general, approximation theory seeks to find an approximation g to f which minimizes $\|f-g\|$ for some appropriate norm $\|\cdot\|:\mathcal{F}\to R^+$. Much of the theory studies whether the relationship

$$(1) \qquad \{\langle f,g\rangle \in \mathcal{F}\times\mathcal{H} \mid \|f-g\| = \inf_{h\in\mathcal{H}} \|f-h\|\}$$

has domain $\mathcal{F}$ (existence) and whether it is a function (unicity).

Modern norms, such as the p-norms, $1 \le p \le \infty$, depend on the entire domain of their argument, so that their computation is necessarily a limiting process, in fact, the foundation of analysis. Our more primitive idea is to base the norm on a subset of the domain small enough so that the approximated function f is summable over it. Thus we want a $P \subset X$ such that $\sum_{x\in P} f(x)$ makes sense, and so we will use the countable analysis of infinite series.

It was discovered early, and is confirmed every day by those untrained numerical analysts, that requiring polynomials to take on given values does not always mean that other values of the polynomial will remain near the given ones. Thus interpolation may not even be good approximation in the uniform sense. Hermite is well known as a pioneer of the theory for improving this situation by also specifying the values of the derivatives as well as function values. Differentiation, like function evaluation, is a linear operator, so that the Hermite conditions constitute specifying the values of linear functionals as a way of making f and g match. Thus we can increase generality considerably but not restrict our theory any by considering the space of linear functionals

$\mathcal{F}^* := \{L \mid L: \mathcal{F} \to K, \text{ linearly}\}$ on $\mathcal{F}$. We choose our subset $\mathcal{L}$ then from $\mathcal{F}^*$ rather than X. Certainly $\mathcal{F}^*$ is no smaller than X if we assume that $\mathcal{F}$ is rich enough to include the constants. Then, for the interpolation condition, we consider approximation in the semi-norm,

$$\|f\|_{\mathcal{L}} := \sqrt{\sum_{L \in \mathcal{L}} |L(f)|^2},$$

but we require, as mentioned, that the approximation be exact. Thus we consider the relation

$$\pi := \{\langle f, g\rangle \in \mathcal{F} \times \mathcal{H} \mid \|f - g\|_{\mathcal{L}} = 0\}. \tag{2}$$

Because all terms of the sum defining the $\mathcal{L}$-norm are non-negative, this amounts to

$$\pi = \{\langle f, g\rangle \mid L(f) = L(g),\ \forall L \in \mathcal{L}\}. \tag{3}$$

Given $\mathcal{F}$, $\mathcal{H}$ and $\mathcal{L}$, we will refer to (3) as the *interpolation relation* and say that $\mathcal{H}$ *interpolates* $\mathcal{F}$ *over* $\mathcal{L}$ in case the relation is a function on $\mathcal{F}$. Certainly π is onto since it is a super-relation of the identity. If π is a function, it is clearly a projection, being linear and idempotent. Thus we call it the *interpolation projection.*

2. Algebra

In the case π is well defined as a function on $\mathcal{F}$, our analytical approximation condition has produced an algebraic object in the form of a map on our function space. In pursuing this direction, we will follow Garrett Birkhoff [1978]. He factors π by considering maps by way of a third space $K^{\mathcal{L}}$. This can be thought of as the set of all value sets that interpolants will take on. It is also a vector space over K in the usual way. Its dimension is influenced by the linear independence of $\mathcal{L}$ as we shall see. The components of its vectors are the values of all the functionals in $\mathcal{L}$ expressed at an element of $\mathcal{F}$. When $\mathcal{H}$ and span $\mathcal{L}$ are comparable, the components may dually be thought of as a functional expressed across a basis of $\mathcal{H}$. Note that elements of $K^{\mathcal{L}}$ need not be linear, i.e. we do not exclude $K^{\mathcal{L}} \not\subset (\mathcal{F}^*)^*$. Indeed $K^{\mathcal{L}} \subset (\mathcal{F}^*)^*$ only makes sense if $\mathcal{L}$ is a vector space which it need not be, particularly if it is finite.

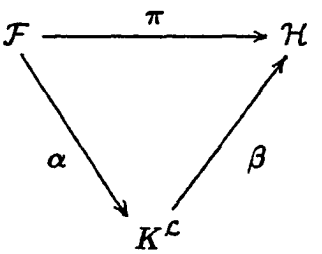
Nonetheless, we can use the canonical evaluation map between $\mathcal{F}$ and $(\mathcal{F}^*)^*$ to get an *interpolation map*

$$\alpha: \mathcal{F} \to K^{\mathcal{L}}$$

$$\text{by} \qquad (\alpha f)(L) := L(f). \tag{4}$$

In this context α is linear, by virtue of inheriting the property from L. Birkhoff [1978] defines any right inverse β of α to be an *interpolation scheme*. Its range, $\beta(K^{\mathcal{L}})$, determines the approximating subspace $\mathcal{H}$, and β picks out an interpolant there given any value set in $K^{\mathcal{L}}$. This gives us the following diagram

(5)



That this diagram commutes, i.e. $\pi = \beta\alpha$, follows from

$$\begin{aligned} g &= (\beta\alpha)f \\ \alpha g &= (\alpha\beta\alpha)f \\ &= \alpha f. \end{aligned}$$

But from the definition (4) of α, $\alpha f = \alpha g$ means

$$\begin{aligned} \forall L \in \mathcal{L} \quad (\alpha f)(L) &= (\alpha g)(L) \\ L(f) &= L(g) \end{aligned}$$

which is the definition (3) of $g = \pi f$.

It is clear from (4) that α (and therefore β) is linear if and only if the functionals of $\mathcal{L}$ are linear. Birkhoff [1978] defines interpolation in the general case where $\mathcal{L}$ is arbitrary in $K^{\mathcal{F}}$ but discusses only the linear case, as we shall also.

Eliminating the limiting process defining approximation in (1) in favor of the vanishing of a discrete norm defining interpolation in (2) simplifies our problem to the extent that important questions can be answered algebraically. However, as remarked earlier, we may have lost ground in that interpolants may differ greatly in the continuum norms of (1). Thus, most practical studies include an error estimate bounding $\|f - g\|$ and seek interpolation functional sets $\mathcal{L}$ for which this bound is well behaved. This is primarily an analytic process and will be pursued here only as it is of algebraic use. In interpolation applications there is usually an analytic step in the proof of existence or uniqueness, but

we will be concerned primarily with the relationships between $\mathcal{F}$, $\mathcal{H}$, $\mathcal{F}^*$ and $\mathcal{L}$ which lead to existence and uniqueness of interpolants for arbitrary elements of the solution space. As noted, $\mathcal{F}$ is usually specified by the application and $\mathcal{H}$ is determined by practical computational considerations. In some cases it is necessary to reduce $\mathcal{F}$, say, by imposing a growth condition on its elements, in order to obtain uniqueness. However, in most cases the interpolator's skill lies in choosing $\mathcal{L}$ so that the interpolation relation is constrained enough to be a function but not so much as to be undefined on some elements of $\mathcal{F}$. These are algebraic considerations which will focus on the maps α, β, and π and other maps related to them.

3. Linear Algebra

By cleaving to linearity we make available the tools of linear representation and all of the power of linear algebra that this brings. We introduce here standard terminology in a fairly standard notation for approximation and analysis. We use the standard textbook formalism as presented in [18], except that we generalize to infinite dimensional spaces in two standard analytical ways. See for instance [14].

Thus for a subset $\mathcal{U}$ of a vector space $\mathcal{H}$ we will consider the *span* of $\mathcal{U}$ to be the space *generated* by all linear combinations of elements of $\mathcal{U}$:

$$\text{span}\,\mathcal{U} := \{v = \sum_{i=1}^{m} a_i u_i \mid m \in N^+,\ \mathbf{a} \in K^m,\ \mathbf{u} \in \mathcal{U}^m\}. \tag{6}$$

in which case we say that v is *represented* over $\mathcal{U}$ by $\mathbf{a}$ and $\mathbf{u}$. Here N^+ is the positive integers, the $\sum$ is from the linear structure of $\mathcal{H}$ and, as usual in an applied context, the boldface notation indicates an m-tuple of elements from the appropriate set:

$$\mathbf{u} := \langle u_1, \ldots, u_m \rangle^T, \qquad u_i \in \mathcal{U}. \tag{7}$$

The transpose superscript indicates that we will be using the standard row, column and matrix notation of linear algebra. Boldface capitals will be matrices

$$\mathbf{B} := [b_{ij}]_{i,j=1}^{m,p} \in \mathcal{U}^{m \times p}$$

of elements usually from K, $\mathcal{H}$ or $\mathcal{H}^*$. In the latter cases the argument $x \in X$ or $h \in \mathcal{H}$ will not necessarily share the boldface. For example

$$\mathbf{u}(x) := \langle u_1(x), \ldots, u_m(x) \rangle^T \in \mathcal{H}^m.$$

Juxtaposition will denote the standard matrix products in $\mathcal{U}^m$ and $\mathcal{U}^{m\times p}$ with the dot product written as

$$\mathbf{u}\cdot\mathbf{v} := \langle u_1,\ldots,u_m\rangle\langle v_1,\ldots,v_m\rangle^T.$$

The symbol u of (7) may be a literal element of K so that

$$\mathbf{0} := \langle 0,0,\ldots,0\rangle^T.$$

With the convention that upper and lower case numerals are the same, we can also use $\mathbf{0}$ for the zero matrix.

The Kronecker delta function is the characteristic function of the identity relation on any set $\mathcal{U}$, i.e.

$$\delta{:}\mathcal{U}^2 \to K$$

$$\text{by} \qquad \delta_{ij} := \delta(i,j) := \begin{cases} 1, & \text{if } i=j; \\ 0, & \text{otherwise.} \end{cases}$$

The standard orthonormal basis for K^m is

$$\{\mathbf{e}_i := \langle \delta_{i1},\delta_{i2},\ldots,\delta_{im}\rangle^T\}_{i=1}^m$$

and the $m\times m$ identity matrix is

$$\mathbf{I}_m := [\mathbf{e}_1,\ldots,\mathbf{e}_m].$$

As the span of a set extends it to the minimal larger object which is a vector space, the idea of linear independence reduces a set to the minimal one which will generate the space. Thus we eliminate linear combinations of the most primitive elements by defining a set to be *linearly independent* if no non-trivial linear combination is zero:

$$\mathcal{U} \text{ is linearly independent} := \forall v\in \operatorname{span}\mathcal{U} \quad v=0 \iff \mathbf{a}=\mathbf{0},$$

where $\mathbf{a}$ represents v over $\mathcal{U}$. If $\mathcal{H}=\operatorname{span}\mathcal{B}$ and $\mathcal{B}$ is linearly independent, we say that $\mathcal{B}$ is a *basis* for $\mathcal{H}$.

We will also consider an infinite linear representation when K has an absolute value. Then

$$\operatorname{span}_\infty\mathcal{U} := \{v \mid \lim_{m\to\infty}\sum_{i=1}^{m}|a_iu_i(x)| \text{ exists and}$$

$$\lim_{m\to\infty}|v(x)-\sum_{i=1}^{m}a_iu_i(x)|=0,\ \mathbf{a}\in K^{N^+},\ \mathbf{u}\in\mathcal{U}^{N^+}\},$$

where the limits are in the sense of uniform convergence on X. Generally we will only consider the special case of the span of a given countable basis $\{u_i\}_{i=1}^{\infty}$. In fact, sorting out what spaces can be represented by which countable bases is a major concern of functional analysis. The crucial thing that concerns us here, however, is the uniqueness of representation. It is not surprising then that the formulation of linear independence is such as to guarantee uniqueness of representation.

THEOREM: *$\mathcal{B}$ is a basis for $\mathcal{H}$ if and only if*

$$\forall v \in \mathcal{H}\ \exists!\ m \in N^+,\ \exists!\ \mathbf{a} \in K^m,\ \exists!\ \mathbf{u} \in \mathcal{B}^m \quad v = \mathbf{a} \cdot \mathbf{u}.$$

In the case $\mathcal{H} = \text{span}_{\infty}\{u_i\}_{i=1}^{\infty}$, we also have unique representation when $\{u_i\}$ is linearly independent, which, in this context is rendered

$$\forall x \in X \quad \sum_{i=1}^{\infty} a_i u_i(x) = 0 \iff a_i = 0 \quad \forall i \in N^+.$$

The proof in each case appeals to the linearity of $\sum$ to show that any two representations must have the same elements.

The size of a basis is a fundamental characteristic of a vector space, particularly when it is finite. Thus call a vector space *finitely generated* if it is the span of a finite basis. The uniqueness of representation then assures us that all bases have the same size. Define this common size to be the *dimension* of the space. A space which can not be finitely generated is said to have *infinite dimension.*

4. Duality

We have defined the vector space of linear functionals $\mathcal{H}^*$. This space is *dual* to the parent space $\mathcal{H}$ in that their bases can be related to each other by a combination of the evaluation map and the Kronecker delta known as *biorthonormality*. If $\mathcal{U}$ is a basis for $\mathcal{H}$, define

$$\nu: \mathcal{H} \to \mathcal{H}^*$$

$$\text{by} \quad (\nu u)(v) := \delta(u, v) \quad u, v \in \mathcal{U}$$

and extend this map linearly by

$$\nu \sum a_i u_i := \sum a_i \nu u_i$$

to get a linear transformation from a vector space to its dual. Dually, if $\mathcal{L}$ is a basis of $\mathcal{H}^*$, define

$$\nu^*: \mathcal{H}^* \to \mathcal{H}$$

$$\text{by} \quad L(\nu^* F) := \delta(L, F) \quad L, F \in \mathcal{L}$$

and extend linearly, also. The fundamental property of these maps is given by

THEOREM: a) *$\mathcal{U}$ is a basis for $\mathcal{H} \Longrightarrow \nu(\mathcal{U})$ is a basis for $\mathcal{H}^*$.*

b) *$\mathcal{L}$ is a basis for $\mathcal{H}^* \Longrightarrow \nu^*(\mathcal{L})$ is a basis for $\mathcal{H}$.*

COROLLARY: *For finite dimensional spaces, ν and ν^* are bijective.*

The proof of a) verifies in a straightforward way the linear independence of $\nu(\mathcal{U})$ due to the biorthogonality relationship whereby νu_j extracts the j^{th} coefficient in the representation over $\mathcal{U}$ of a vector. For this reason it is called the j^{th} *coordinate functional.* For infinite representation, elementary facts of absolute convergence are needed. That $\nu(\mathcal{U})$ is complete in $\mathcal{H}^*$ follows from the representation

(9). $$L = \sum_{j \in I} L(u_j) \nu u_j.$$

To prove part b) we apply a dual argument, first showing that $\nu^*(\mathcal{L})$ is linearly independent. Thus let a linear combination of the $\nu^* L_i$'s be zero,

$$\sum_{i \in I} c_i \nu^* L_i = 0.$$

Then

$$\forall j \in I \quad L_j\Big(\sum_i c_i \nu^* L_i\Big) = 0$$

$$\sum_i c_i L_j(\nu^* L_i) = 0$$

$$\sum_i c_i \delta_{ij} = 0$$

$$c_j = 0.$$

Lastly we show $\mathcal{H} = \operatorname{span} \nu^*(\mathcal{L})$ by considering a representation dual to (9). Thus we claim that for every $v \in \mathcal{H}$

(11) $$v = \sum_{j \in I} L_j(v) \nu^* L_j.$$

To show identity in (11) we propose that every linear functional $L = \sum a_i L_i$ in $\mathcal{H}^*$ takes the same value on the right hand side of (11) as it does on v. Since this would include all coordinate functionals, unique representation establishes the identity. Thus apply L to the right hand side of (11) to get

$$\begin{aligned}\Big(\sum_{i\in I} a_i L_i\Big)\Big(\sum_{j\in I} L_j(v)\nu^* L_j\Big) &= \sum_{i,j\in I} a_i L_j(v) L_i(\nu^* L_j)\\ &= \sum_{i,j} a_i L_j(v)\delta_{ij}\\ &= \sum_{i} a_i L_i(v)\\ &= L(v).\end{aligned}$$

The dual basis elements $l_j := \nu^* L_j$ are known as the *cardinal* basis functions because of the distinguished way that they represent elements of $\mathcal{H}$ in (11). From the definition of ν^* we then have that this distinguishing property is biorthonormality:

$$L_i(l_j) = \delta_{ij}, \quad i,j \in I.$$

The representation (11) is called the *Lagrange representation* since Lagrange was the first to use it for univariate pointwise polynomial interpolation . The cardinal basis is therefore also known as the *Lagrange basis*.

In the case that $\mathcal{L}$ is finite, there is a "dual" Gram-Schmidt procedure for construction of the corresponding cardinal set in $\mathcal{H}$. See [14]. The part played by the inner product in the standard Gram-Schmidt procedure is taken here by a "duality product" given by evaluation of a linear functional on a function. We will see later that such a "product" can, in fact, be developed into a genuine inner product on the function space and that the standard Gram-schmidt procedure can deliver biorthonormality from it. The "dual" procedure uses induction on the number m of functionals in $\mathcal{L}$, with the existence of an element, $h_m \in \mathcal{H}$ for which $L_1, \ldots, L_{m-1}$ vanish but $L_m(h_m) \neq 0$, guaranteed by linear independence of $\{L_1, \ldots, L_m\}$. Otherwise, the Gram-Schmidt linear combination gives the dependancy.

When a finite $\mathcal{L}$ is a basis for $\mathcal{H}^*$, both $\mathcal{H}$ and $\mathcal{H}^*$ have the same finite dimension m and the cardinal basis $\{l_i\}$ of $\mathcal{H}$ is uniquely determined by the biorthonormality

relationship. For if $\{u_i\}$ were some other basis which were also biorthonormal to $\mathcal{L}$, we could express it in terms of $\{l_i\}$ to get

$$u_k = \sum_{i=1}^{m} b_i l_i \tag{12}$$

$$\begin{aligned} \delta_{jk} = L_j(u_k) &= L_j\Big(\sum b_i l_i\Big) \\ &= \sum b_i L_j(l_i) \\ &= \sum b_i \delta_{ij} \\ &= b_j. \end{aligned}$$

So only b_k is non-zero and so the representation (12) collapses to $u_k = l_k$. This uniqueness is certainly to be expected from the fact that $\{l_i\}$ and $\{L_i\}$ are bases of isomorphic finite dimensional vector spaces.

For a computational basis $\{u_i\}$, let the transition matrix to the Lagrange basis $\{l_i\}$ be C so that

$$l_j = \sum_{k=1}^{m} c_{kj} u_k$$

$$\iff \qquad l_j = \mathbf{u}^T \mathbf{C} \mathbf{e}_j. \tag{13}$$

Then

$$\begin{aligned} (\nu u_i)(l_j) &= \sum_k c_{kj} (\nu u_i)(u_k) \\ &= \sum_k c_{kj} \delta_{ik} \\ (\nu u_i)(\nu^* L_j) &= c_{ij}. \end{aligned}$$

For reasons which will soon become apparent, we call this matrix

$$\mathbf{C} = [(\nu u_i)(\nu^* L_j)]_{i,j=1}^{m,n} \tag{14}$$

the *dual Gramian* of $\mathcal{L}$ and $\mathcal{U}$.

5. Dual Interpolation Map

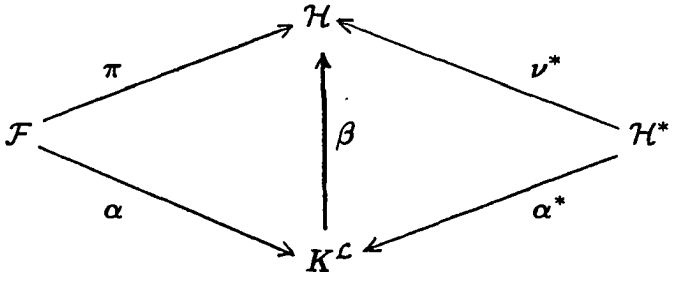
Consider the induced evaluation map dual to α,

$$\alpha^* : \mathcal{F}^* \to K^{\mathcal{L}}$$

$$\text{by} \qquad (\alpha^* F)(L) := F(\nu^* L) \quad F \in \mathcal{F}^*,\ L \in \mathcal{L}. \tag{15}$$

It assigns to a generalized interpolation condition element, the value set which that element expresses over the basis $\{l_i\}$ dual to $\mathcal{L}$. We call this linear transformation the *dual interpolation map*. It completes the assimilation of the dual space $\mathcal{H}^*$ into our map diagram (5) to produce the extended diagram

(16)



6. Singularity

Recall that in §2 we defined β so that $\alpha|_{\mathcal{H}}\beta$ is the identity map on $\alpha(\mathcal{F})$ and we proved there that $\pi = \beta\alpha$. It is worth noting that the interpolation map α carries the non-injective property of the interpolation relation in the case when π is a projection and we have unique interpolation. This is because $\mathcal{H}$ is proper in $\mathcal{F}$ so that there are $f \in \mathcal{F}$ such that $f \notin \mathcal{H}$. But then $g = \pi f \in \mathcal{H}$ has been shown to be such that $\alpha(g - f) = 0$ while $g - f \neq 0$. A map which carries a non-zero element to zero is said to be *singular*. For linear maps the condition is equivalent to non-injectivity and non-invertibility. Thus a necessary condition for interpolation is that α be singular as a linear transformation on $\mathcal{F}$.

In our expanded diagram (16) we show the maps ν^* and α^* restricted to the computational subspaces $\mathcal{H}$ and $\mathcal{H}^*$. We will see that non-singularity is the issue for the restricted maps and that, as might be expected, this makes the diagram commute so that $\nu^* = \beta\alpha^*$, and guarantees interpolation of $\mathcal{F}$ by $\mathcal{H}$ over $\mathcal{L}$.

7. Fundamental Theorem of Finite Interpolation

For the case of finite $\mathcal{L} := \{L_i\}_{i=1}^n$, the space of value sets $K^{\mathcal{L}}$ can be identified with K^n by $\gamma \leftrightarrow \langle \gamma(L_1), \gamma(L_2), \ldots, \gamma(L_n) \rangle$ for each $\gamma \in K^{\mathcal{L}}$. Now, as anticipated, all spaces are finite dimensional and the interpolation map may be explicitly written out from (4) as

$$\alpha f = \mathbf{L}(f) := \langle L_1(f), \ldots, L_n(f) \rangle. \tag{17}$$

Here we abuse notation by making $\mathbf{L}$ a vector, not a matrix. Applying (17) to the basis $\{u_i\}_{i=1}^m$ of $\mathcal{H}$ shows that $\mathbf{G} = [L_i(u_j)]_{i,j=1}^{n,m}$ is the resulting representation of $\alpha|_{\mathcal{H}}$. This matrix is called the *Gramian* of $\mathcal{L}$ over $\mathcal{U}$. For a given value set $\mathbf{a}$, it produces the linear system $\mathbf{Gb} = \mathbf{a}$ which expresses the interpolation problem to be solved for $\mathbf{b}$ in order to find the interpolant $g := \mathbf{b} \cdot \mathbf{u}$ in terms of the basis $\{u_i\}$.

We now finally have all the tools at hand to discuss interpolation and the fundamental theorem concerning the solvability of the interpolation problem. Recall from the definition (3) that, $\mathcal{H}$ interpolates $\mathcal{F}$ over $\mathcal{L}$, means that for every element $f \in \mathcal{F}$ there is a unique element $g \in \mathcal{H}$ such that $L(f) = L(g)$ for every $L \in \mathcal{L}$. From (17) we see that this amounts to existence of the interpolation scheme β onto $\mathcal{H}$ acting as an inverse for α to select the interpolant g from $\mathbf{L}(f)$. Existence of an inverse for α by linearity amounts to non-singularity of $\alpha|_{\mathcal{H}}$. By definition this means no non-zero element of $\mathcal{H}$ is mapped onto the zero value set $\mathbf{0}$, i.e. no non-zero interpolant interpolates the zero function. It is clear that if there were such, this would spoil uniqueness of interpolation, since then any multiple of the the non-zero interpolant could be added to any other interpolating function to get a different function of the approximating subspace which also interpolated.

Also from linear algebra we can deduce another equivalent condition, namely that $n = m$ and that $\det \mathbf{G} \neq 0$. Then we can get the interpolation scheme directly as

$$\beta\mathbf{a} = \mathbf{u}^T\mathbf{G}^{-1}\mathbf{a} \qquad \mathbf{a} \in K^n. \tag{18}$$

If we can exhibit a set $\{l_i\}$ which is biorthonormal to $\mathcal{L}$, then we can solve the interpolation problem explicitly in terms of that set by the representation (11) which renders (18) as (with $\boldsymbol{\ell}$ for boldface l)

$$\beta\mathbf{a} = \boldsymbol{\ell}\cdot\mathbf{a} \qquad \mathbf{a} \in K^n, \tag{19}$$

since for the basis $\{l_i\}$ the Gramian is $\mathbf{I}_n$. We can then write the interpolation projection as

$$\pi f = \sum_{i=1}^{n} L_i(f) l_i$$
$$(\pi f)(x) = \mathbf{L}(f)\cdot\boldsymbol{\ell}(x).$$

From equation (19), with $\mathbf{e}_j$ for $\mathbf{a}$, we have

$$\beta\mathbf{e}_j = \boldsymbol{\ell}\cdot\mathbf{e}_j = l_j.$$

Combining this with the same instance of (18) gives

$$l_j = \mathbf{u}^T \mathbf{G}^{-1} \mathbf{e}_j.$$

In effect we have solved the interpolation problem for the value set which has the Lagrange basis functions for interpolants. Comparing the above equation with (13), we conclude from the uniqueness of representation of l_j with respect to $\{u_i\}$

THEOREM:

$$\mathbf{C} = \mathbf{G}^{-1}. \tag{20}$$

Recall that $\mathbf{C}$ is the dual Gramian (14).

The dual interpolation map is explicitly, from (15)

$$\alpha^* F = \langle F(l_1), \ldots, F(l_n) \rangle \qquad F \in \mathcal{F}^*. \tag{21}$$

Representing α^* with respect to the basis $\{\nu u_i\}$ of $\mathcal{H}^*$ so that $F = \mathbf{b} \cdot \nu \mathbf{u}$, we have

$$\begin{aligned} F(l_j) &= \sum_{i=1}^{n} b_i (\nu u_i)(l_j) \\ &= \sum_i b_i (\nu u_i)(\nu^* l_j) \\ &= \sum_i b_i c_{ij}. \end{aligned}$$

Thus the dual Gramian earns its name by serving to represent, with its transpose, the dual interpolation map. Then we know from (20) that non-singularity of $\alpha^*|_{\mathcal{H}^*}$ will produce an inverse for the Gramian. This gives us still another equivalence for interpolation. We also may conclude thereby that β is non-singular and so from (19) and (21) that for any $L_i \in \mathcal{L}$,

$$\begin{aligned} (\beta\alpha^*) L_i &= \beta \langle L_i(l_1), \ldots, L_i(l_n) \rangle \\ &= \beta \mathbf{e}_i \\ &= \ell \cdot \mathbf{e}_i = l_i, \end{aligned}$$

which shows that $\nu^* = \beta\alpha^*$ since they have the same action on $\{L_i\}$ which forms a basis since ν^* is bijective. Conversely, independence of $\{L_i\}$ ensures bijectivity of ν^* and so non-singularity of α^* which provides a final fifth equivalent condition for interpolation.

This and the definitions of α, β and ν^* with the help of (20) imply,

$$\begin{aligned} l_i &= \beta\alpha^* L_i \\ \mathbf{e}_i = \alpha l_i &= \alpha^* L_i \\ L_i &= (\alpha^*)^{-1}\mathbf{e}_i \\ L_i &= (\nu\mathbf{u})^T(\mathbf{C}^T)^{-1}\mathbf{e}_i \\ &= (\nu\mathbf{u})^T\mathbf{G}^T\mathbf{e}_i, \end{aligned}$$

so that the transition matrix from $\{\nu u_i\}$ to $\{L_i\}$ is the transpose of the Gramian.

Introducing the notation, $F(\mathbf{u}) := \langle F(u_i), \ldots, F(u_m) \rangle$, we can now collect all of these characterizations to write our fundamental theorem as follows.

THEOREM: *The following are equivalent*

a) $\mathcal{H}$ *interpolates* $\mathcal{F}$ *over* $\mathcal{L}$

b) $\alpha|_{\mathcal{H}}$ *is non-singular, i.e.,*

$\forall g \in \mathcal{H} \quad \mathbf{L}(g) = \mathbf{0} \Longrightarrow g = 0$

c) $\det[L_i(u_j)] \neq 0$

d) $\exists \{l_i\} \subset \mathcal{F} \qquad L_i(l_j) = \delta_{ij} \quad i,j = 1, \ldots, n.$

e) $\alpha^*|_{\mathcal{H}^*}$ *is non-singular, i.e.,*

$\forall F \in \mathcal{H}^* \quad F(\ell) = \mathbf{0} \Longrightarrow F = 0$

f) $\operatorname{span}\{L_i|_{\mathcal{H}^*}\} = \mathcal{H}^*$

We should also note that while a) specifies that f and πf should match at all elements of $\mathcal{L}$, linear representation assures us that they will, in fact, match at all elements of span $\mathcal{L}$ and so, by virtue of f), over $\mathcal{H}^*$ as well. Another weakening of hypotheses which is recovered is the uniqueness of interpolants guaranteed in a). If we require that the interpolants merely exist for the limited collection of value sets $\{\mathbf{e}_i\}_{i=1}^n$, then they will be the required $\{l_i\}_{i=1}^n$ of d). Then the machinery of biorthonormality gives the uniqueness and representation for all value sets. Thus a linear system is invertible for all elements of a space, provided it is invertible on a basis.

An equivalent way to express the non-singularity condition b), is to say that an element of $\mathcal{H}$ is uniquely determined by its values over $\mathcal{L}$. Thus if two functions agree

on $\mathcal{L}$, they must be identical. This useful *identity* condition often shows that two different formulations of an element of $\mathcal{H}$, in fact, denote the same object. The dual approach for e) yields the more obvious fact that a linear functional is uniquely specified by its values on a basis of its domain.

There are two ways that interpolation can fail. We can have α singular and $m \leq n$ which implies the existence of a non-zero $\mathbf{a} \in K^m$ such that $\mathbf{Ga} = \mathbf{0}$. Then $h := \mathbf{a} \cdot \mathbf{u}$ is an example violating b) and we have a manifold, $\{g + ah \mid a \in K\}$, of solutions to the interpolation problem, rather than just one. Conversely, if $m \geq n$, we have that α^* is singular and there is a non-zero $\mathbf{b} \in K^n$ such that $\mathbf{Cb} = \mathbf{0}$. Then $F := \mathbf{b} \cdot \nu\mathbf{u}$ violates e) and there is some $f \in \mathcal{F}$ such that $F(f) \neq 0$. This f has no interpolant g because g would have to be represented over $\{l_i\}$ by (19). But $F(g) = 0$ since F annihilates span $\{\nu^* L_i\}$. Being non-zero on f it cannot match f with any such g.

8. Interpolation Inner Product

If our field is rich enough to have a conjugation automorphism $\overline{\cdot}: K \to K$ for which $a\overline{a} = |a|^2$, we can define a bilinear functional

$$\varphi: \mathcal{H}^2 \to K$$

$$\text{by} \quad \varphi(u,v) := \sum_{i=1}^{n} L_i(u)\overline{L_i(v)} = \mathbf{L}(u) \cdot \overline{\mathbf{L}}(v)$$

which generates the interpolation semi-norm by

$$\|u\|_{\mathcal{L}} = \sqrt{\varphi(u,u)}.$$

Now non-singularity condition b) implies that φ is positive definite so that it is an inner product and, therefore, that $\| \cdot \|_{\mathcal{L}}$ is a proper norm. Under these conditions the Gram-Schmidt procedure enables us to construct a basis which is orthonormal with respect to this inner product. That is

$$\mathcal{H} = \text{span}\,\{v_i\}_{i=1}^{n}$$

and

$$\begin{aligned} \delta_{ij} &= \varphi(v_i, v_j) \\ &= \mathbf{L}(v_i) \cdot \overline{\mathbf{L}}(v_j). \end{aligned}$$

Let G be the Gramian of $\mathcal{L}$ with respect to $\{v_i\}$, i.e.

$$\mathbf{G} = [L_i(v_j)]_{i,j=1}^n.$$

Then

$$\begin{aligned} \mathbf{G}\mathbf{G}^H &= \left[\sum_k L_k(v_i)\overline{L_k(v_j)}\right] \\ &= [\delta_{ij}] = \mathbf{I}_n \end{aligned} \tag{22}$$

where the superscript H indicates the conjugate transpose or *Hermitian* transform of a matrix. The identity (22) characterizes *unitary* matrices for which $\mathbf{G}^{-1} = \mathbf{G}^H$. Then we also have

$$\begin{aligned} \mathbf{I}_n &= \mathbf{G}^H\mathbf{G} \\ [\delta_{ij}] &= \left[\sum_k \overline{L_i(v_k)}L_j(v_k)\right]. \end{aligned} \tag{23}$$

Then if we define

$$l_i(x) := \sum_k \overline{L_i(v_k)}v_k(x),$$

by (23) we have

$$L_j(l_i) = \sum_k \overline{L_i(v_k)}L_j(v_k) = \delta_{ij},$$

so that $\{l_i\}$ is biorthonormal to $\mathcal{L}$ as desired.

9. Notes

The material in this chapter has been extrapolated primarily from the works three authors, P. J. Davis's book [1975], the three papers by H. C. Thacher, et al [1958] [1960a] [1960b], and C. de Boor's textbook [1978] and his Mathathematics Research Center monograph [1982]. My notation follows Davis and Thacher although de Boor is more contemporary and complete. There seems to be no agreed upon uniform notation. As my scope is wider, including some of the material from Birkhoff [1978], I have attempted to be consistent within textbook style guidelines. De Boor is the only author to introduce an acronymic predicate to describe the Linear Interpolation Problem. He writes LIP($\mathcal{H}, \mathcal{L}$) to mean "given $f \in \mathcal{F}$, find $g \in \mathcal{H}$ such that $L(f) = L(g), \quad \forall L \in \mathcal{L}$."

Then he speaks of a LIP$(\mathcal{H}, \mathcal{L})$ as *correct* in case $\mathcal{H}$ interpolates $\mathcal{F}$ over $\mathcal{L}$ in my sense. Usually the space $\mathcal{F}$ contains functions to be approximated. For algebraic purposes, however, $\mathcal{F}$ need only be large enough to express all possible value sets $K^{\mathcal{L}}$. If $\mathcal{L}$ is relatively small (often finite), it is customary to replace "for all $f \in \mathcal{F}$" with "for all value sets equinumerous with $\mathcal{L}$." Then the overlying solution space is properly implicit and de Boor's notation conveys all the information.

The maps α, β, etc. are after Birkhoff [1978], but none of the authors uses the dual interpolation map α^*. In my development it brings in the dual Gramian and shows that the transpose Gramian is the transition matrix to the functional set. It also helps prove that interpolation is equivalent to the linear independence of the functional set $\mathcal{L}$ (Fundamental Theorem (e)), a result which Davis demonstrates directly via linear equations theory. The construction of the Lagrange basis by way of the interpolation inner product and Gram-Schmidt procedure does not appear in the literature I have been able to survey, although I have heard data analysts speak of orthonormalizing a computational basis with respect to a data set.

CHAPTER II

Applications to Finite Multivariate Lagrange Interpolation on Regular Grids

1. Non-standard Example: Linear Ordinary Differential Initial Value Problems

Interpolation texts invariably begin with real univariate polynomials of degree $n-1$ interpolating arbitrary value sets in R^n over n distinct point evaluations. We have made an effort to be general from the start, so let us demonstrate the power of the fundamental theorem by examining a more advanced example. For some $n \in N^+$ let

$$\mathcal{F} = C^n(a,b) := \{f \in R^{(a,b)} \mid f^{(n)}(x) \text{ is continuous on } (a,b)\}$$

where $(a,b) := \{x \in R \mid a < x < b\}$, $f^{(n)}$ is the n^{th} derivative of f, and $f^{(0)} := f$. For some $\mathbf{p}(x) \in (C^0(a,b))^n$ define

$$\mathcal{H} = \{f \in \mathcal{F} \mid f^{(n)} = \sum_{k=1}^{n} p_k f^{(k-1)}\}.$$

Because differentiation is linear, $\mathcal{H}$ is a subspace of $\mathcal{F}$. Complete the interpolation triple by defining, for some $x_0 \in (a,b)$

$$\mathcal{L} = \{L_k \in \mathcal{F}^* \mid L_k(f) = f^{(k-1)}(x_0), \quad k = 1,\ldots,n\}.$$

Now the basic existence and uniqueness theorem of general linear ordinary differential initial value problems can be stated succinctly as $\mathcal{H}$ interpolates $\mathcal{F}$ over $\mathcal{L}$. By our finite interpolation theorem, we know that $\mathcal{H}$ is an n-dimensional space and that there must therefore be n linearly independent solutions to the differential equation, and that the "general" solution is an element of their span involving n arbitrary constants. Furthermore, we have the five equivalent characterizations b) through f). Thus only the zero solution is initially at complete rest. Also, the determinant of $[u_i^{(j-1)}(x_0)]$ is not zero. Since x_0 was arbitrary, this generalizes to the statement that the *Wronskian* $W(x) := \det[u_i^{(j-1)}(x)]$ is non-zero throughout (a,b) for any linearly independent set

of solutions $\{u_i\}$ spanning $\mathcal{H}$. This fact is usually proved analytically, by showing that $W(x)$ satisfies a differential equation that implies it has an exponential form, and so has strictly one sign. To get this last from our result, we must appeal to the Intermediate Value Theorem and the continuity of $W(x)$.

Dually, we know that $\{L_i\}$ is linearly independent and a basis for $\mathcal{H}^*$. This is a remarkable assertion: any linear functional on the space of solutions of this differential equation can be represented as a linear combination of some point evaluation over the first n (zero through $n-1$) derivatives of its argument. Thus, as the Reisz Representation Theorem shows that any bounded linear functional on the space of square integrable functions is essentially integration, we have that any linear functional on the solution space of a linear differential equation is essentially differentiation. Conversely, any linearly independent set of n such linear combinations determines a unique solution to the initial value problem. This gives us the capability of imposing a whole matrix of initial conditions at a point,

$$\mathbf{AL}(f) = \mathbf{b},$$

provided **A** is non-singular.

Lastly, we have two methods for constructing a distinguished set $\{l_i\}$ of linearly independent solutions to the differential equation such that, given any initial values $\mathbf{b} \in R^n$, the solution satisfying them can be expressed directly as $g = \mathbf{b} \cdot \ell$.

This approach is a kind of inversion of the usual method, as in Householder [1953], of Taylor expansion to obtain the error. It has been extended by W. J. Gordon [1971] to include functional sets from boundary value problems and multivariate partial differential equations. Note that $\mathcal{H}$ is a space of polynomials in case the differential equation is $f^{(n)} = 0$. Thus we can consider polynomial interpolation to be a special case of this more general scheme. Then the approximation to the larger space C^n results in the residual being a solution to the homogeneous equation. The Green's function gives an integral formula for the Sard kernel error estimates on interpolation and quadrature formulas.

2. Multivatiate Polynomial Spaces

By multivariate spaces we mean subspaces of K^{R^n}. Multivariate polynomials are best described in terms of sets of multi-indices. These are sets of tuples of the nonnegative integers, N_0, usually defined by some constraint on the elements of the tuples. We will

use lower case Greek letters for multi-indices and the corresponding upper case Greek letter for a set of such multi-indices. A superscript on the set will indicate its dimension, i.e., the length of the tuples. Subscripts on the multi-index will indicate its components. Thus

$$\kappa = \langle \kappa_1, \kappa_2, \ldots, \kappa_n \rangle \in \mathrm{K}^n \subset N_0^n.$$

The modulus of a multi-index will actually be its 1-norm as an element of R^n so that $|\kappa| := \sum_1^n \kappa_i$. For $\mathbf{x} \in R^n$ let $\mathbf{x}^\kappa := \prod_{i=1}^n x_i^{\kappa_i}$. Then $\mathbf{x} \mapsto \mathbf{x}^\kappa$ defines a multivariate monomial function. These serve as the basis for the space of all multivariate polynomials,

$$\mathcal{P}^n := \mathcal{P}(R^n) := \mathrm{span}\,\{\mathbf{x}^\kappa \mid \kappa \in N_0^n\}$$

Multi-index sets we will frequently use are those of maximum order m, $\{0, \ldots, m\}^n$, and those of total order m,

$$\Lambda_m^n := \{\lambda \in \{0, \ldots, m\}^n \mid |\lambda| \leq m\}.$$

It is easy to see that the cardinality of this set is given by $\#\Lambda_m^n = \binom{n+m}{m}$. The barycentric multi-index set (whose name will become apparent later) is defined by

$$\Gamma_m^n := \{\gamma \in \{0, \ldots, m\}^n \mid |\gamma| = m\}.$$

Elementary properties of the binomial coefficient show that $\#\Gamma_m^{n+1} = \binom{n+m}{m}$ also. Corresponding to these sets we define polynomial spaces of limited degree. These are the polynomials of *maximum degree* m,

$$\mathcal{P}_m^n := \mathcal{P}_m(R^n) := \mathrm{span}\,\{\mathbf{x}^\kappa \mid \kappa \in \{0, \ldots, m\}^n\},$$

and the polynomials of *total degree* m,

$$\mathcal{T}_m^n := \mathcal{T}_m(R^n) := \mathrm{span}\,\{\mathbf{x}^\lambda \mid \lambda \in \Lambda_m^n\}.$$

Their dimensions are determined by the size of the multi-index sets for their exponents. These are respectively $(m+1)^n$ and $\binom{n+m}{m}$. That the monomials are linearly independent is an elementary theorem of multivariate polynomial algebra.

3. Differentiation and Point Evaluation Functionals

Although there are a great variety of linear functionals, the most commonly used in applications are those obtained by evaluating a derivative of a function at a point of its

domain. In multivariate analysis, the derivatives are partial and, to simplify notation, are specified by multi-indices. Thus the differentiation operator is a map

$$D^{\lambda} : C^{(|\lambda|)}(R^n) \to C^{(|\lambda|-1)}(R^n)$$

defined as

$$D^{\lambda}(f(\mathbf{x})) := \frac{\partial^{|\lambda|} f(\mathbf{x})}{\partial^{\lambda_1} x_1 \cdots \partial^{\lambda_n} x_n}.$$

where $\partial^0 f := f$, $\partial^1 := \partial$ and $\partial^0 x_i$ is null. For a point $\mathbf{z} \in R^n$, define the *evaluation functional* by $E_{\mathbf{z}}(f) := f(\mathbf{z})$. Then the λ-order *Hermite* functional at $\mathbf{z}$ evaluated at f is

$$E_{\mathbf{z}}^{\lambda}(f) := E_{\mathbf{z}} D^{\lambda}(f) = \left. \frac{\partial^{|\lambda|} f(\mathbf{x})}{\partial^{\lambda_1} x_1 \cdots \partial^{\lambda_n} x_n} \right|_{\mathbf{x}=\mathbf{z}}.$$

Now given a *node set* P of points of R^n, and a multi-index set $\Lambda \subset N_0^n$, we can form the set of linear functionals

$$\mathcal{L}_P^{\Lambda} := \{E_{\mathbf{z}}^{\lambda} \mid \lambda \in \Lambda, \mathbf{z} \in P\},$$

and speak of the Λ-order *Hermite interpolation* problem over P. Here we are seeking a finite dimensional subspace of $\mathcal{P}^n$ which will interpolate differentiable functions over $\mathcal{L}_P^{\Lambda}$. In the special case where $\Lambda = \{0\}$ the functionals are restricted to point evaluation and we have the simplest and most commonly encountered case of *Lagrange interpolation.*

4. Univariate Polynomial Lagrange Interpolation

In the univariate case ($n = 1$) the fact that m-degree polynomials interpolate arbitrary values over any set of $m + 1$ of points is well known. The commonest proof uses the elementary theorem of algebra that a polynomial of degree m has at most m roots. Therefore if it vanishes on the node set it must be the zero polynomial. This proves interpolation via non-singularity. We can also note that the Gramian has the Vandermonde determinant which is constructively non-zero. The most useful method, however, is to exhibit the Lagrange basis which not only proves interpolation, but also provides a well conditioned algorithm for computing it. For $P = \{z_i\}_{i=0}^m$ we have

$$l_j(x) = \prod_{\substack{i=0 \\ i \neq j}}^{m} \frac{x - z_i}{z_j - z_i} \qquad j = 0, \ldots, m.$$

5. Multivariate Polynomial Lagrange Interpolation

The well behaved nature of the univariate polynomials, whereby the elements have zero sets strictly smaller than the dimension of the space, is called the *Haar* property and, as noted, ensures interpolation. Higher dimensonal spaces, however, are not Haar. For instance the quadratic $x_1^2 + x_2^2 - 1$ vanishes at all points on the unit circle in R^2 and therefore has an uncountable zero set. In general it is quite difficult to prove that a given set does not lie in the implicitly specified surface $\{\mathbf{x} \mid p(\mathbf{x}) = 0\}$ for *all* multivariate polynomials p in the space. The Gramian matrix becomes huge and unwieldy for even moderate values of m and n although it has been attacked successfully in some notable cases (Schoenberg and Whitney [1953], Karlin, et al [1966] [1968]). The most fruitful efforts have focused on determining conditions on the node set which permit the construction of a Lagrange basis. We now apply this approach to the two most general polynomial spaces.

5.1 Maximum Degree Lagrange Interpolation. This space is characterized easily as an instance of the more general *tensor pròduct* spaces as defined, for example in Cheney [1986] or de Boor [1982]. The multivariate space is constructed from a formal product of elements of univariate spaces, and the Lagrange basis in the product space over a product of functionals is the product of the univariate Lagrange bases. Symbolically

$$\mathcal{P}_m^n \cong \mathcal{P}_m^1 \otimes \mathcal{P}_m^1 \otimes \cdots \otimes \mathcal{P}_m^1,$$

where the equivalence symbol denotes vector space isomorphism. The natural product node set for this space is a rectangular grid,

$$P = \prod_{i=1}^{n} \{z_{ij}\}_{j=0}^{m} = \{\langle z_{1\kappa_1}, \ldots, z_{n\kappa_n}\rangle \mid \kappa \in \{0, \ldots, m\}^n\}.$$

The above double subscripted n-tuple in z will be the meaning of $\mathbf{z}_\kappa$. As noted, the Lagrange basis is a product of component Lagrange polynomials,

$$l_\kappa(\mathbf{x}) = \prod_{i=1}^{n} l_{\kappa_i}(x_i),$$

where

$$l_{\kappa_i}(x) = \prod_{\substack{j=0 \\ j \neq \kappa_i}}^{m} \frac{x - z_{ij}}{z_{i\kappa_i} - z_{ij}}.$$

Although obviously useful for properly placed data, this kind of node set is not well suited for scattered (say naturally occurring) data, or for patching together "elements" of piecewise polynomial function spaces as used for spline approximation and the finite element method. The most effective general method for interpolating scattered data in the plane is triangulation of the node set followed by *blending* of cubic splines (Nelson [1983]). The blending, a term introduced by Coons [1967], matches derivatives at triangle boundaries to ensure a smooth, although piecewise, interpolant. The fundamental construct here is a triangular node set with Hermite interpolation functionals. In n dimensions this generalizes to a simplicial node set with similar functionals. Accordingly we next consider the total degree space for which simplices are natural node sets.

5.2 Total Degree Lagrange Interpolation (TDLI). Principal Lattices. The prototypical simplex is the right unit simplex with its square vertex at the origin,

$$U_1^n := \{\mathbf{0}\} \cup \{\mathbf{e}_i\}_{i=1}^n.$$

Any general n-simplex, $S^n = \{\mathbf{v}_j\}_{j=0}^n$ can be obtained from U_1^n by the affine transform,

$$\mathbf{v} = \mathbf{V}_t\mathbf{x} + \mathbf{v}_0,$$

where $\mathbf{V}_t$ is the *translated vertex* matrix,

$$\mathbf{V}_t := [v_{ij} - v_{i0}]_{i,j=1}^n,$$

and $\mathbf{v}_0$ is the translation from the origin. Of course not just any matrix in $\mathcal{M}^{n\times n}$ will do, for we do not want the vertices to coalesce or degenerate onto a lower dimensional manifold. Thus we need $\mathbf{V}_t$ to represent a non-singular transform onto R^n. This condition is ensured by *affine independence* of the $\mathbf{v}_j$, by which we mean, $\forall \mathbf{c} \in R^{n+1}$

$$\sum_{j=0}^n c_j\mathbf{v}_j = \mathbf{0} \text{ and } \sum_{j=0}^n c_j = 0 \quad \Longrightarrow \quad \mathbf{c} = \mathbf{0}.$$

This in turn is equivalent to the existence of an *affine representation* for every point of R^n in terms of the $\mathbf{v}_j$. That is, $\exists \mathbf{b} \in R^{n+1}$ for which

$$\mathbf{x} = \sum_{j=0}^n b_j\mathbf{v}_j, \qquad \sum_{j=0}^n b_j = 1. \tag{1}$$

This amounts to a bijective mapping from R^n onto the unit slant hyperplane in R^{n+1},

$$\mu : R^n \to H_1^{n+1} := \{\mathbf{b} \in R^{n+1} \mid \sum_{j=1}^{n} b_j = 1\}.$$

The b_j are called the *barycentric* coordinates of $\mathbf{x}$ with respect to S^n. The matrix corresponding to μ^{-1} is the vertex matrix, $\mathbf{V} := [v_{ij}]_{i=1,j=0}^{n}$. We must not think thereby that μ or its inverse is linear. In fact the range of μ is a linear manifold rather than a vector space. Since μ is a bijection we have $\forall \mathbf{x} \in R^n \; \exists! \, \mathbf{b} \in R^{n+1}$ such that $\mathbf{x} = \mathbf{V}\mathbf{b}$. However to compute the barycentric coordinates we need to solve the larger system (1) which is $\mathbf{V}_0\mathbf{b} = \mathbf{x}_0$, where $\mathbf{x}_0 = \langle 1, x_1, \ldots, x_n \rangle$ and,

$$\mathbf{V}_0 = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ v_{10} & v_{11} & \cdots & v_{1n} \\ v_{20} & v_{21} & \cdots & v_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ v_{n0} & v_{n1} & \cdots & v_{nn} \end{bmatrix} = \begin{bmatrix} \mathbf{1}^T \\ \mathbf{V} \end{bmatrix}.$$

Thus we need $\mathbf{V}_0 \in \mathcal{M}^{n+1 \times n+1}$ to be invertible at least on the manifold of points with constant first coordinate $\mathbf{x}_0 \in \langle 1, 0, \ldots, 0 \rangle + \{0\} \times R^n$.

For the simplex U_1^n, the vertex matrix is $\mathbf{V}_1 = [\mathbf{0}\, \mathbf{I}_n]$ which implies a barycentric matrix

$$\mathbf{V}_{1,0}^{-1} = \begin{bmatrix} 1 & -1 & \cdots & -1 \\ 0 & & & \\ \vdots & & \mathbf{I}_n & \\ 0 & & & \end{bmatrix}.$$

This makes the barycentric coordinates the same as the cartesian ones in addition to the barycentric combination,

$$b_j = \begin{cases} x_j, & j \geq 1; \\ x_0 := 1 - \sum_{i=1}^{n} x_i, & j = 0. \end{cases}$$

In this case the barycentric mapping μ has the canonical form,

$$\langle x_1, \ldots, x_n \rangle \leftrightarrow \langle 1 - \textstyle\sum_i x_i, x_1, \ldots, x_n \rangle.$$

Consider a similar correspondence between Λ_m^n and Γ_m^{n+1} by

$$\langle \lambda_1, \ldots, \lambda_n \rangle \leftrightarrow \langle m - |\lambda|, \lambda_1, \ldots, \lambda_n \rangle. \tag{2}$$

Then $\frac{1}{m}\Gamma_m^{n+1}$ are the barycentric coordinates of the node set $U_m^n := \frac{1}{m}\Lambda_m^n$ which is just the intersection of a regular grid of spacing $\frac{1}{m}$ with the convex hull of the right unit simplex U_1^n. We call U_m^n the right unit m^{th} order principal lattice. Generally, the m^{th} order principal lattice of an arbitrary simplex, $S^n = \{\mathbf{v}_j\}_{j=0}^n$, are those points whose barycentric coordinates are $\frac{1}{m}\Gamma_m^{n+1}$. In vertex notation,

$$S_m^n := \mu^{-1}(\frac{1}{m}\Gamma_m^{n+1}) = \mathbf{V}_t(U_m^n) + \mathbf{v}_0,$$

which is a non-singular affine transform of the right unit m^{th} order principal lattice. The top diagram of Figure 1 shows the case for $n = m = 3$. The nodes are labeled by their barycentric indices. The lower number on each index gives its lexicographic order.

We note finally that an affine transform is a first total degree polynomial function. Since polynomial composition multiplies degree, we have that the total degree polynomial space is invariant under affine transformation of the variable space. Thus if a polynomial $p(\mathbf{x})$ assumes given values on the node set U_m^n, then $p(\mathbf{V}_t^{-1}(\mathbf{x} - \mathbf{v}_0))$ will be a polynomial of the same total degree which assumes the same values on the corresponding nodes in S_m^n. Therefore we need only study the interpolation problem on U_m^n in order to solve it for any simplex. This normalizing approach is also computationally preferable as any ill condition of the system can be isolated in the matrix $\mathbf{V}_t$. Lagrange basis polynomials on U_m^n will involve quantities depending only on n and m, not on the size or relative spacing of vertices in the simplex.

Much of this section is the relatively elementary linear algebra of vector geometry as presented, for instance, in Nering [1963]. We reproduce it here to consolidate notation and because it is not used in the earlier treatment of this problem in the literature. Wading back and forth through the affine transformations in the general simplicial setting is extremely tedious and time consuming. Normalization simplifies things considerably.

5.3 TDLI. Normalized Simplicial Lagrange Basis. Following the previous constructions, form a product of one degree terms in each variable involving the difference between the variable and a corresponding node component z_{ij} where $\mathbf{z}_j$ are the nodes of U_m^n. Since these components are multiples of $\frac{1}{m}$, we can normalize to form

$$\prod_{j \neq k} \frac{m x_i - \kappa_{ij}}{\kappa_{ik} - \kappa_{ij}} \tag{3}$$

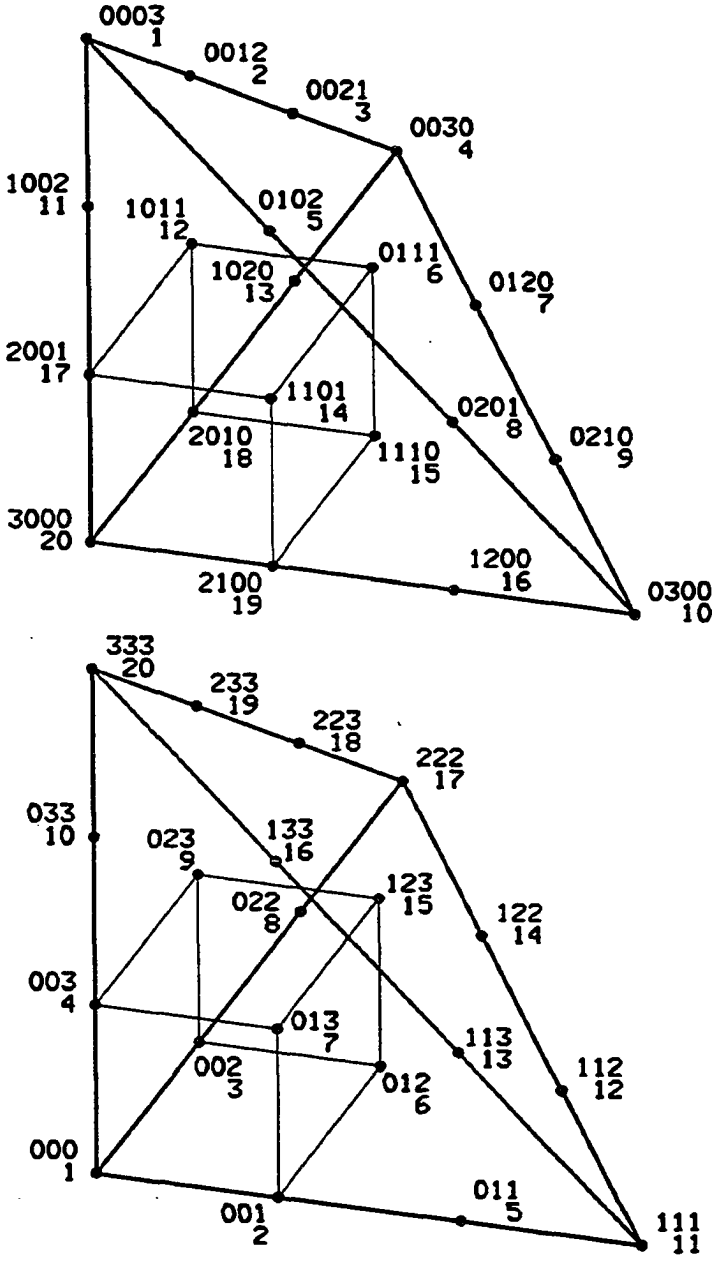


Figure 1. Two descriptions of U_3^3, the right normalized third order, three dimensional, principal simplicial lattice at the origin. In the upper diagram each node is labeled by the appropriate element of Γ_3^4, so that the last three digits are the normalized cartesian coordinates. The number just below each barycentric multi-index gives its position in ascending lexicographic order. The nodes of the lower simplex are labeled by the elements of Φ_3^3. The vertex on the x_i-axis is labeled with the index iii. Nodes lying interior to a subcell of dimension q of the simplex are labeled with indices having $q+1$ distinct components chosen from the integers labeling the vertices on the boundary of the subcell. When more than one node occupies the same subcell, the components are repeated according to the proximity of the corresponding boundary vertex. The numbers below the indices give ascending lexicographic order. That this geometric placement corresponds to reverse lexicographic order of the barycentric indices is readily checked.

where the κ_{ij} are integers in the range 0 to m. All such products form the maximum degree polynomial. So we eliminate node components κ_{ij} which make the denominator of (3) negative as well as zero. This follows a triangular pattern appropriate to the simplex. Further triangularize by using barycentric coordinates and so include x_0 in (3). Since barycentric and cartesian coordinates are affine combinations of each other, a total degree polynomial in one system is total degree in the other.

From these motivations, and a good bit of trial and error (see Appendix I), we arrive at the following form. With each node, by the canonical barycentric mapping (2), associate a $\gamma \in \Gamma_m^{n+1}$. Then the cardinal basis function at that node is

$$l_\gamma(\mathbf{x}) := \prod_{i=0}^{n} \binom{mx_i}{\gamma_i} \tag{4}$$

where $\binom{a}{p}$ is the conventional binomial coefficient for real a and integer $p \geq 0$, and where $\binom{a}{0} = 1$ for all a. Then each of the $n+1$ factors is a univariate polynomial of degree γ_i in x_i including $i = 0$. So $\deg l_\gamma = \sum_0^n \gamma_i = |\gamma| = m$, i.e., $l_\gamma \in \mathcal{T}_m^n$ for all γ. It remains to show biorthonormality between $\{l_\gamma \mid \gamma \in \Gamma_m^{n+1}\}$ and U_m^n to ensure that we have a Lagrange basis.

That l_γ is unity on its "own" node is easy. The cartesian coordinates of a node are $\frac{1}{m}\lambda$ for some element of Λ_m^n. From (2) get the barycentric version $\frac{1}{m}\gamma$ so that $\gamma = \langle m - |\lambda|, \lambda_1, \ldots, \lambda_n \rangle$. Then

$$l_\gamma(\frac{1}{m}\lambda) = \prod_{i=0}^{n} \binom{m(\gamma_i/m)}{\gamma_i} = \prod_i \binom{\gamma_i}{\gamma_i} = \prod 1 = 1.$$

To show that l_γ vanishes at all other nodes with barycentric coordinates $\frac{1}{m}\gamma'$ for $\gamma' \neq \gamma$, we need a property of the barycentric index set.

LEMMA: *Let γ', γ be distinct elements of Γ_m^{n+1}. Then $\exists k \in \{0, 1, \ldots, n\}$ such that $\gamma'_k < \gamma_k$.*

If not, i.e., $\forall k\ \gamma'_k \geq \gamma_k$, then $\gamma' \neq \gamma \Longrightarrow \exists j, \quad \gamma'_j > \gamma_j$, and so $\sum_0^n \gamma'_i > \sum_0^n \gamma_i$. Then either $|\gamma'| \neq m$ or $|\gamma| \neq m$ so not both are in Γ_m^{n+1} $\gg\ll$.

Now let γ', γ be as in the lemma. Then the numerator of $\binom{\gamma'_k}{\gamma_k}$ is

$$\prod_{j=0}^{\gamma_k - 1} (\gamma'_k - j).$$

Since $0 \leq \gamma_k' \leq \gamma_k - 1$, j must "pass" γ_k' so that one of the factors must be zero. Then the entire product is zero and we have proved the second part of biorthonormality.

Formula (4) is about as lean as it can be and still identify nodes, variable, dimension, degree, coordinates and form. We can follow Nicolaides [1972] and Chung and Yao [1977] and make the barycentric coordinates for a general simplex explicit. They write

$$(5) \qquad l_{\mathbf{z}_j}(\mathbf{x}) = \prod_{\substack{i=0 \\ b_i(\mathbf{z}_j)>0}}^{n} \prod_{k=0}^{mb_i(\mathbf{z}_j)-1} \frac{mb_i(\mathbf{x}) - k}{mb_i(\mathbf{z}_j) - k},$$

where each principal lattice node $\mathbf{z}_j$ is such that $mb_i(\mathbf{z}_j) \in \Gamma_m^{n+1}$. Their proofs of this form are necessarily more involved but use essentially the same logic, in particular, the multi-index lemma.

5.4 Computing the TDLI Basis Functions. Besides being more concise and better conditioned, formula (4) indicates possibilities for improvement. The binomial coefficient notation implicitly eliminates the unit factors where $\gamma_i = 0$. This case is checked explicitly in the Chung and Yao formula. Thus there are $n + 1$ computations of barycentric coordinates. Then for those which are positive, 1 to m factors must be computed, but the total number of factors is always exactly m. Since all nodes on the surface of the simplex have some zero barycentric coordinates, and such nodes are the majority in many practical applications, there will be several null factors in the formula. However, we have normalized away the affine transform by working with the easiest simplex. We should no longer need it to determine the form of the polynomial. Ideally it should be possible to eliminate n from (4) and instead associate with each node an m-tuple of pairs of integers $\langle \langle \phi_k, \psi_k \rangle \rangle_{k=1}^{m}$ such that

$$(6) \qquad l_{\mathbf{z}_\phi}(\mathbf{x}) = \prod_{k=1}^{m} (m x_{\phi_k} - \psi_k + 1)/\psi_k$$

where the ϕ's and ψ's are cross referencing index maps from the node identifier which we here anticipate will be ϕ since it will have to be an m length multi-index. It should be noted that (6) is not a different formula from (4). Indeed the factors occurring in each will be identical. The purpose of (6) is to express the formula with a minimum of operations. Then the evaluation of the polynomial will involve no redundant computations.

Since the size of the node set is $\binom{n+m}{n}$, we need a multi-index set of this many m-tuples. To serve their purpose of indexing the barycentric coordinates, the ϕ_k should be in the range 0 to n. This leads to the natural choice

$$\Phi_n^m = \{\langle \phi_1, \ldots, \phi_m \rangle \mid 0 \leq \phi_1 \leq \phi_2 \ldots \leq \phi_m \leq n\}.$$

It is an elementary exercise to show $\#\Phi_n^m = \binom{n+m}{n}$.

Olmsted [1986] has shown how to apply geometric considerations in order to associate an element of Φ_n^m with every node of U_m^n in such a way that the proper variable x_{ϕ_k} is selected in (6). The example $n = m = 3$ is shown in the lower part of Figure 1. As with the upper diagram, the lower number of the pair at each node gives the lexicographic order of its index. From ϕ we can compute the correct value of ψ_k by

$$\psi_k = \sum_{i=1}^{k} \delta(\phi_i, \phi_k) \tag{7}$$

which counts the number of prior occurrances of ϕ_k in ϕ.

The details of the correspondence between U_m^n and Φ_n^m are obscure, but simple to compute once worked out. Essentially we have three maps:

$$U_m^n \leftrightarrow \Gamma_m^{n+1} \leftrightarrow \Phi_n^m \rightarrow \{1, \ldots, m\}^m.$$

The first is (2)

$$\mathbf{z}_\lambda = \frac{1}{m}\langle \lambda_1, \ldots, \lambda_n \rangle \leftrightarrow \langle m - |\lambda|, \lambda_1, \ldots, \lambda_n \rangle = \gamma. \tag{8}$$

The third map (neither injective nor surjective) determines the constant term and scale ψ_k of the 1-degree factor by (7). The middle correspondence is constructed by setting the ϕ_k to be the subscripts of the non-zero components γ_j of γ repeated γ_j times. Once this correspondence is computed for various n and m it becomes apparent that it amounts to reverse lexicographic order between the two sets. I have no proof for this fact but it is surely true by the heuristic law of small numbers. Figure 1, of course, confirms the case $n = m = 3$.

Since algorithms to lexicographically list the various multi-index sets are easy to code and fast to run, we now have an effective procedure to compute interpolants to

principal simplicial lattices. If the values to be interpolated are $\{a_\lambda \mid \lambda \in \Lambda_m^n\}$ we need to evaluate

$$p(\mathbf{x}) = \sum_{\lambda \in \Lambda_m^n} a_\lambda l_{\mathbf{s}_\lambda}(\mathbf{x}).$$

We use Γ_m^{n+1} in lexicographic order as a base index set. From it by the inverse barycentric map we order the value set $a_\lambda \leftrightarrow \gamma$ by (8). Reverse lexicographic order selects the proper correspondence to Φ_n^m and, from it, we can finally use (7) to order the ψ's which we need to compute the factors in (6). Since there are no conditional arithmetic steps, this algorithm is optimal up to the speed of indirect addressing in the computer. A completely efficient coding will eliminate the $+1$ operation by actually storing yet another array of the $\psi_k - 1$'s. This fine detail concerning time-space tradeoffs is enough computer science for a thesis in mathematics.

If we follow the above procedure, but at the evaluation step instead write out an expression for the polynomial $l_\phi(\mathbf{x})$, we get a table as shown in Appendix I for the case $n = 4$, $m = 5$. Note that there we have not listed ϕ and l_ϕ for every γ, but have taken advantage of natural groupings, according to values of the ψ's, such that each group has the same form of l_ϕ, and corresponds to a subset of nodes all lying within a subcell (subsimplex) of the simplex. When these subcells are listed by dimension, we see an orderly relationship between the geometry of the principal lattice and the form of l_ϕ as determined from the indices. A kind of lexicographic ordering again appears amongst the groups of the barycentric multi-indices ϕ. How this might be related to reverse lexicographic ordering of Γ_m^{n+1}, however, is not clear to me.

CHAPTER III

Applications to Multivariate Newton-Hermite Interpolation on Subtriangular Grids

1. Splines and Finite Elements

Spline interpolation is a subject that Minsky could appreciate. Neither Hermite nor Laguerre worked in it. Most references are within living, or even computer, memory. Yet as P. J. Davis remarked in 1964 (quoted in Schoenberg [1982]), it "contains the delicious paradox of Prokofieff's Classical Symphony. It seems as though it might have been written several centuries ago, but of course it could not have been." The level of the mathematics is elementary. Hermite or Laguerre would have been able to comprehend it with no special effort, yet they would have wondered why. It takes the very applied aspects of modern engineering and computer graphics to motivate splines. Hermite or Laguerre probably would have done the job by hand.

Splining approximates with much more versatile function spaces than the polynomial ones. The essential idea is that the behavior of the interpolant far from the node set is immaterial to the process of approximation. The TDLI is defined everywhere in R^n. But at distances thousands of times the diameter of the node set, its computation would be very ill-conditioned if meaningful at all, since any polynomial is unbounded at infinity. We therefore bind the approximants to the node set by including piecewise polynomials in our basis. Then the computation of the basis function will depend on the proximity of the argument to the nodes. The atomic basis functions are of the form $(x - z_\kappa)_+$ where z_κ is a node coordinate, and the $+$ subscript indicates a zero value for $x < z_\kappa$.

In the one dimensional case, we arrive at an interpolant which is a polynomial of fixed degree in each internode interval, $z_\kappa \leq x \leq z_{\kappa+1}$. At end points, adjacent polynomials match in value. Thus a polygonal line through $\{\langle z_\kappa, a_\kappa \rangle\}$ is a first degree spline. For higher degree piecewise polynomial interpolation we need more constraints in order to obtain uniqueness. Lacking data, we turn to smoothness conditions, which take the form of constraints on the values of derivatives at the nodes. Thus a cubic, having four terms, can be made to interpolate two nodes, both in function values and in first derivative values. This is a Hermite interpolation problem, known as *osculatory*

interpolation, leading to a linear system in the appropriate functionals. When these systems are solved simultaneously on each internode interval, along with conditions on the values of derivatives at the end nodes, a unique piecewise cubic spline function is obtained. By construction it is continuously differentiable. In this sense three is the "natural" degree for smooth interpolation between nodes on a line. Lower degree interpolants are not smooth. Higher degree piecewise polynomials are underdetermined, except by higher order smoothness constraints.

In dimensions greater than one, the situation is immensely complicated. Internode intervals now correspond to regions of n-space, so determining basis functions on such regions requires precise knowledge of how to calculate the boundaries of the region. For simplicity we assume the boundaries are hyperplanes so that the regions are determined by systems of linear inequalities. I.e., they are polyhedral complexes. Just as scattered points in the plane can be triangulated, such regions can be subdivided into simplices. The central problem then amounts to a Hermite interpolation problem on a simplicial node set. As with the univariate case, for any given order of smoothness, there is a "natural" degree polynomial which solves the interpolation problem for data values at given nodes, as well as matching derivative values across inter-element boundaries. Since, however, the boundaries are continua, it is necessary to discretize them by placing additional nodes on them to get a set of the proper size and configuration. Partial and directional derivatives are then constrained at these locations. The major talent in this subject lies in choosing the proper nodes and derivatives so as to get linear independance among them, and so unique interpolation.

A grand master of this art is A. Ženíšek. He has worked out the details for arbitrary smoothness in two dimensions and first and second order smoothness in three dimensions. For the last mentioned case, the resulting monstrous interpolating polynomial is of seventeenth $(2^3 2 + 1)$ total degree and so has $\binom{17+3}{3} = 1140$ terms. There are six kinds of differentiation functionals of up to eighth (2^3) order evaluated at sixty-nine nodes on the three dimensional tetrahedron. To give a flavor for this kind of thing, we paraphrase Ženíšek [1973] on m^{th} order continuous bivariate piecewise polynomials on a triangulated region of the plane.

Let a triangle be $P_1P_2P_3$ with barycenter P_0. Divide the sides P_iP_j into $k+1$ equal parts by k points $\{Q_r^{(k)} \mid r = 1,\ldots,3k,\ k = 1,\ldots,m\}$. Now form the functional sets,

$$\mathcal{L}_v = \{E_{P_i}^{\lambda} \mid \lambda \in \Lambda_{2m}^2,\ i = 1,2,3\},$$
$$\mathcal{L}_0 = \{E_{P_0}^{\lambda} \mid \lambda \in \Lambda_{m-2}^2\},$$
$$\mathcal{L}_s = \{E_{Q_r^{(k)}} \frac{\partial^k}{\partial n^k} \mid \begin{matrix} k = 1,\ldots,m,\ r = 1,\ldots,3k, \\ \frac{\partial}{\partial n} := \text{derivative normal to the sides} \end{matrix}\}.$$

Simple enumerations give $\#\mathcal{L}_v = 3\binom{2m+2}{2}$, $\#\mathcal{L}_0 = \binom{m}{2}$, and $\#\mathcal{L}_s = 3\binom{m+1}{2}$. Adding these and simplifying shows that, miraculously, the cardinality of the functional set $\mathcal{L} := \mathcal{L}_v \cup \mathcal{L}_0 \cup \mathcal{L}_s$ is $\binom{2+(4m+1)}{2}$ so that, if it is independent, it will uniquely determine a bivariate interpolant of total degree $4m+1$.

To show independence, Ženíšek [1970] kills two birds with one stone. First, to get an error bound for the approximation, he proves a theorem which bounds, on the closed triangle, all derivatives of order up to $4m$ (including order zero) of any function f which is annihilated by the set $\mathcal{L}$, in terms of bounds on the $4m+2^{nd}$-order derivatives of f. When applied to the residual $f - \pi f$, it gives an error estimate in terms of the max over the triangle of the derivatives of f of order *one greater* than the total degree of the interpolant. In particular it requires

$$|f| \leq c(\max\{|D^{\lambda} f(\mathbf{x})| \mid |\lambda| = 4m+2, \quad \mathbf{x} \in P_1P_2P_3\})$$

where c is a constant independent of f. Since any polynomial of total degree $4m+1$ has zero for derivative of order $4m+2$, this shows that such a polynomial, if annihilated by $\mathcal{L}$, will be identically zero on the triangle. Since the dimension of the root set of a non-zero polynomial is strictly less than the dimension of its domain, we conclude that any polynomial annihilated by $\mathcal{L}$ must be the zero polynomial. This shows independence, and so interpolation, by non-singularity.

2. Computing Hermite Interpolants.

Ženíšek's analytic proof gives error estimates, but existence is shown indirectly. Our earlier formulas are not Hermite, nor do they or other "natural" lattices which interpolate over total degree polynomial spaces have a flexible enough number of geometric/differential data conditions. It should not be thought thereby that interpolable node configurations are rare. As Cheney [1986] notes, the probability of multivariate polynomial interpolation on a random node set is unity, since the root set has zero measure

being of strictly lesser dimension than the object space. Our focus shifts, therefore, to the computational problem of evaluating the interpolant in terms of the functional set.

2.1 Newton Formulation Newton divided differences are presented in every numerical analysis text (Householder [1953]). The original method is the essence of triangularization. Successive differences of adjacent data are listed in a table which terminates at a single highest order value. The coefficients representing the interpolant over the basis of monomials of differences, $\{\prod_{i<j}(x - z_i) \mid j = 0, \ldots m\}$ are selected from the table. The selection order symmetrizes the procedure so that the resulting expression is invariant under permutation of the data.

The method has two great beauties. The first is the recursive structure which enables an algorithm to accomodate new data without reinitializing. De Boor [1982] notes that this amounts to selecting new bases, $\{u_i'\}$ and $\{L_i'\}$, for $\mathcal{H}$ and $\mathcal{H}^*$ whose transition matrices, $\mathbf{L}^{-1}$ and $\mathbf{U}^{-1}$, to the given bases $\{u_i\}$ and$\{L_i\}$ come from a triangular decomposition (or factorization) of the Gramian $\mathbf{G} = \mathbf{LU}$.[1] Back substitution in the triangular system of equations gives the recursive scheme for successive computation of u_{i+1}'. Davis [1975] shows how (with a Gram-Schmidt/Gauss flavor) the recursion proves such a decomposition is always possible when $\det \mathbf{C} \neq 0$, and that the resulting new bases are biorthogonal, $L_i'(u_j') = \delta_{ij}$. Cheney [1986] abstracts further to show how to merge the solutions of the interpolation problem on two disjoint node sets, into a solution of the interpolation on the union of the node sets.

2.2 Univariate Newton-Hermite Representation. The second great beauty of Newton's formulation is its ability to express derivative values. The coefficients in the Lagrange representation, $L_i(f)$ are as simple as they can be. To get more power, we complicate by considering a functional which now depends on the first i functionals. It is usually denoted by listing the functionals inside square brackets. Furthermore, to simplify notation, we assume evaluation functionals and suppress the 'E' functional name, as well as move the argument f of the functional to the left. This is consistent with Newton's classical notation. De Boor [1978], ever vigilant for mathematically consistent (and therefore non-standard) notation, does not make the last compromise but puts the

[1] (Here we perpetrate, now obversely, abuse of notation on the overused symbols L and U)

argument on the right. Having gone part way, we prefer the familiar confusion, and so define the divided difference functional by the following recursion on the zero indexed nondescending *sequence* $z_0, \ldots, z_n$ of nodes.

$$f[z_0] := f(z_0).$$

If the node values are distinct,

$$f[z_0, z_1, \ldots, z_i] := \frac{f[z_0, \ldots, z_{i-1}] - f[z_1, \ldots, z_i]}{z_0 - z_i}, \tag{10}$$

else for repeated nodes,

$$f[z_0, z_0, \ldots, z_0] := \frac{1}{i!} f^{(i)}(z_0). \tag{11}$$

Another feature made explicit by de Boor's more precise notation, is that this recursion goes both ways. Nodes may be added on the left as well as the right

$$f[z_{-1}, z_0, \ldots, z_{i-1}] := \frac{f[z_{-1}, z_0, \ldots, z_{i-2}] - f[z_0, \ldots, z_{i-1}]}{z_{-1} - z_{i-1}},$$

so that the recursion direction may be alternated at will. Thus any expression of the form

$$f[z_0, z_1, \ldots, z_k, z_k, \ldots, z_k, z_{k+1}, \ldots, z_i]$$

may be acquired via the recursion (10) from $f[z_k, \ldots, z_k]$ which in turn is evaluated with (11). Then the Newton representation is

$$\pi f = \sum_{i=0}^{n} f[z_0, \ldots, z_i] \prod_{j<i} (x - z_j).$$

There are many fine properties of the Newton divided difference but, as can be anticipated by form, and proved by mean value theorems, the one which makes the above definition consistent, and introduces derivative values, is

$$\lim_{\forall k, z_k \to z_0} f[z_0, z_1, \ldots, z_i] = \frac{1}{i!} f^{(i)}(z_0).$$

Using this fact, we can write a solution to the general Hermite interpolation problem which interpolates to order m_i at each point y_i. The fundamental multi-index set, which we call *Newton-Hermite* is

$$I_{\mathbf{m}}^{n} := \{\langle i, j \rangle \mid \begin{matrix} i = 0, \ldots, n \\ j = 0, \ldots, m_i \end{matrix}\},$$

ordered lexicographically. Its size is $\sum_{i=0}^{n}(m_i+1)$. Define the node sequence $\{z_\iota\}_{\iota\in I_{\mathbf{m}}^n}$ by repeating m_i times each y_i so that $z_{\langle i,j\rangle} := y_i,\ j = 0,\ldots,m_i$. Then the functional set can be written

$$\mathcal{L} = \{L_{\langle i,j\rangle} := E_{y_i}^j \mid \langle i,j\rangle \in I_{\mathbf{m}}^n\},$$

and

$$\mathcal{L} = \{L_{\langle i,j\rangle} := E_{y_i}^j \mid \langle i,j\rangle \in I_{\mathbf{m}}^n\},$$

interand we show that arbitrary functions can be interpolated by $\mathcal{P}^1$ over $\mathcal{L}$ by exhibiting the interpolant in the Newton representation

$$\pi f = \sum_{\iota\in I_{\mathbf{m}}^n} f[z_{\langle 0,0\rangle},\ldots,z_\iota]\prod_{\kappa<\iota}(x - z_\kappa).$$

This method of multiplicity of nodes, also known as *coalescence* or *merging* of the nodes, produces basis functions with positive powers $(x-y_i)^{m_i}$ of the factor at each node where there are derivative values to be matched. For these terms, the divided difference coefficients are computed using derivative values in (11) at the appropriate step of the recursion.

2.3 Specifying Multivariate Hermite Functionals. For simplicity we start with dimension 2. Then a Hermite functional is a directional derivative in the plane. Full information includes direction and order. Direction is specified by a line

$$H = H(a,b,c) := \{\langle x,y\rangle \mid ax + by + c = 0, \quad |a| + |b| \neq 0\}.$$

This redundant (essentially homogeneous) three parameter specification is needed in order to include either direction on the line. Now differentiate by

$$D_{H(a,b,c)}f := -b\frac{\partial f}{\partial x} + a\frac{\partial f}{\partial y}.$$

Then $D_{H(-a,-b,-c)}$ is the derivative in the opposite direction. Lastly let

$$p_{a,b,c}(x,y) := ax + by + c$$

be the first degree polynomial associated with the line $H(a,b,c)$. Then for all x,y

$$D_{H(a,b,c)}p_{a,b,c}(x,y) = -ba + ab = 0.$$

Thus we should include the factor $p_{a,b,c}$ in basis functions for nodes at which we wish to specify a derivative along $H(a,b,c)$. For higher order derivatives, we include the appropriate power of this first degree factor. In this way a Hermite functional at a point in the plane can be identified with a collection of lines through that point.

Gasca and Maeztu [1982] have devised an explicit scheme from this concept. They incorporate additionally that a "natural" node set in the plane tends to be grouped along lines. Thus they have primary and secondary lines but provide for differentiation along either of them. Their base object is

$$S := \{\langle H_i, J_i\rangle\}_{i=0}^{n},$$

a collection of pairs of lines,

$$H_i = \{\langle x,y\rangle \mid H_i = H(a_i,b_i,c_i)\},$$

and sequences of secondary lines,

$$J_i = \{H_{ij} \mid H_{ij} = H(a_{ij},b_{ij},c_{ij}),\ j = 1,\ldots,m_i\}.$$

Amazingly, the only restriction we need is that the secondary (or cross) lines do, in fact, determine a discrete set of nodes by their intersections with the primary ones,

$$\{z_{ij}\} = H_i \cap H_{ij}, \quad \forall\langle i,j\rangle \in I_{\mathbf{m}}^{n}.$$

The definitions of the sequences $i \mapsto H_i$ and $j \mapsto H_{ij}$ do not require the images to be distinct. Indeed, we will want a multiplicity of lines in order to get derivative functionals. Note, however, that the specification of two lines, H, H', in terms of their coefficients a,b,c may differ but the lines will consist of the same points if $\langle a',b',c'\rangle = r\langle a,b,c\rangle$ for some non-zero $r \in R$. The derivatives will, however, differ by the constant factor r. We note the interesting fact that the set S is indexed by the same Newton-Hermite set $I_{\mathbf{m}}^{n}$ as in the univariate case. The dimension of the contents, however, has increased by one. The point set

$$\{z_{ij}\}_{\langle i,j\rangle \in I_{\mathbf{m}}^{n}},$$

may be of smaller cardinality because of multiplicity, i.e., $\langle i,j\rangle \mapsto z_{ij}$ will not be injective in the Hermite case.

Thus we define functionals

$$L_{ij} := (E_{z_{ij}} D^{r_{ij}}_{H_{ij}})(E_{z_{ij}} D^{s_{ij}+t_{ij}}_{H_i}) \tag{12}$$

where

$$D^k_{H(a,b,c)} = (-b\frac{\partial}{\partial x} + a\frac{\partial}{\partial y})^k,$$

and the orders of differentiation are given by

$$\begin{aligned} r_{ij} &:= \sum_{k<i} \delta(H_k, H_i), \\ s_{ij} &:= \#\{k \mid z_{ij} \in H_{ik}, \quad k \le j-1\}, \\ t_{ij} &:= \#\{k \mid z_{ij} \in H_k \neq H_i, \quad k \le i-1\}. \end{aligned}$$

In effect, we differentiate at z_{ij} along H_{ij} with order according to the multiplicities of previous occurrances (in the lexicographic order) of identical primary lines through the node z_{ij}. We differentiate along the primary line with order according to two multiplicities: the number of cross lines (of the i^{th} primary) which pass through z_{ij} which are also prior in order to the particular cross line H_{ij} in question, plus the number of primary lines distinct from and prior to primary line H_i in consideration.

It would seem that there is much ambiguity in this definition. Many configurations of lines could specify the same functional. In fact for the evaluation functional of order zero, any two lines intersecting at the node will do. As long as they are not parallel, their directions are irrelevant. We will see that these ambiguities are "divided out" in the definitions of the next sections which specify the basis functions and their coefficients.

2.4 Bivariate Newton-Hermite Basis Functions. The work having gone before, the definition is easy. Relabel the first degree polynomials associated with each line

$$\begin{aligned} p_i(x,y) &:= a_i x + b_i y + c_i \\ p_{ij}(x,y) &:= a_{ij} x + b_{ij} y + c_{ij}. \end{aligned}$$

Then define the polynomial basis functions $\{u_{ij}(x,y)\}_{\langle i,j\rangle \in I_m^n}$ by

$$u_{ij} := \Big(\prod_{k<i} p_k\Big)\Big(\prod_{k<j} p_{ik}\Big) \tag{13}$$

where i or $j = 0$ implies an empty product with the value unity. Note that (13) gives the total degree of u_{ij} as $i+j$. Thus span $\{u_{ij}\} \subset \mathcal{T}_m^2$, where $m = \max\{i+j \mid \langle i,j\rangle \in I_m^n\}$, but the inclusion is most likely proper. The major result of Gasca and Maeztu [9] is then

THEOREM: *$\{u_{ij}\}$ is a Newton basis with respect to $\{L_{ij}\}$, i.e.,*

$$\langle i,j\rangle < \langle i',j'\rangle \implies L_{ij}(u_{i'j'}) = 0$$
$$\text{but } \forall \langle i,j\rangle \in I_m^n \quad L_{ij}(u_{ij}) \neq 0,$$

where the functionals L_{ij} are as defined in (12). This produces a lower triangular Gramian with non-zero diagonal entries which therefore has non-zero determinant. This demonstrates that span $\{u_{ij}\}$ interpolates arbitrary values over $\{L_{ij}\}$ by (c) of the fundamental theorem.

2.5 Bivariate Newton-Hermite Representation. The final step in solving this interpolation problem is to compute the coefficients in the representation of the interpolant over $\{u_{ij}\}$. The back substitution method amounts to a double recursion in f_{ij}, the generalized divided differences, and $q_{ij}(x,y)$, the partial solution up to condition L_{ij}. Explicitly this is

$$f_{00} := f(z_{00})$$
$$q_{00} := f(z_{00})$$

$$f_{k,l+1} = \frac{L_{k,l+1}(f) - L_{k,l+1}(q_{kl})}{L_{k,l+1}(u_{k,l+1})}, \qquad l < m_k \tag{14}$$
$$f_{k+1,0} = \frac{L_{k+1,0}(f) - L_{k+1,0}(q_{km_k})}{L_{k+1,0}(u_{k+1,0})}$$

$$q_{k,l+1} = q_{kl} + f_{k,l+1}u_{k,l+1}, \qquad l < m_k$$
$$q_{k+1,0} = q_{km_k} + f_{k+1,0}u_{k+1,0}.$$

Now we can represent the partial interpolant,

$$q_{kl} = \sum_{\langle i,j\rangle \leq \langle k,l\rangle} f_{ij} u_{ij},$$

and the final interpolant to f as

$$\pi f = q_{n,m_n} = \sum_{\langle i,j\rangle \in I^n_{\mathbf{m}}} f_{ij} u_{ij}.$$

Note that every interpolation problem solvable by this method must include at least one zero-order evaluation functional to start the recursion. Thus it is immediately apparent that we cannot interpolate over every functional set (e.g. $\mathcal{L} = \{E^1_{z_{00}}\}$) by this method.

2.6 Trivariate Newton-Hermite Representation. Gasca and Maeztu [1982] outline the manner in which this method can be extended to R^3 and higher dimensions. The concept is further triangular nesting. The dimension of the base objects is upped to two and the index set is given three levels. The base set is

$$\{H_i, \{H_{ij}, \{H_{ijk}\}_{k=0}^{p_{ij}}\}_{j=0}^{m_i}\}_{i=0}^{n},$$

where each H_κ is a plane specified by four parameters,

$$H_\kappa = \{\langle x, y, z\rangle \mid a_\kappa x + b_\kappa y + c_\kappa z + d_\kappa = 0\}.$$

The geometric requirement is point intersections over index levels,

$$\{z_{ijk}\} = H_i \cap H_{ij} \cap H_{ijk}.$$

The triple index set is

$$I^n_{\mathbf{m},\mathbf{p}} = \{\langle i,j,k\rangle \mid i = 0,\ldots,n,\ j = 0,\ldots,m_i,\ k = 0,\ldots,p_{ij}\},$$

and the basis functions are

$$u_{ijk} = \prod_{u<i} p_u \prod_{v<j} p_{uv} \prod_{w<k} p_{uvw},$$

where $p_\kappa \leftrightarrow H_\kappa$ according to the obvious extension of the two dimensional case.

Functional specification involves six conditions on the various planes and their intersections, which latter determine the direction of differentiation functionals. The formula for expressing L_{ijk} contains such a welter of compounded sub and superscripts that (12) looks relatively tame. I refer the curious reader to Gasca and Maeztu [1982]. Their formula shows in general how we would expect $n(n+1)/2$ conditions on intersecting hyperplanes in dimension n. For the final Newton recursion step, the formulas of (14) are generalized and extended to advance the indices within each level as well as to advance from primary through secondary to tertiary levels. Thus, generically,

$$\begin{aligned} f_{k,l,h+1} &= F(q_{klh}), \qquad h < p_{kl} \\ f_{k,l+1,0} &= G(q_{klp_{kl}}), \qquad l < m_k \\ f_{k+1,0,0} &= H(q_{km_kp_{km_k}}). \end{aligned}$$

This formula has broken through to the triple subscript level so we will skip the simpler expressions for the q_{ijk} recursions and call a halt to any further eye strain.

Overall it appears to me that the method of Gasca and Maeztu [1982] cries out for normalization and computerization of the index manipulations. Then it would be feasible to code machine evaluations of basis functions and the Newton recursion for problems as complex as the Ženíšek three dimensional spline element.

2.7 Lagrange versus Newton-Hermite in the Plane. For this thesis, however, we will examine simpler examples. In fact, one of the simplest Hermite examples in two dimensions illustrates nicely how the Lagrange approach fails but Newton saves the game. Thus a minimal multi-point problem involving two dimensional differentiation has the functional set

$$\mathcal{L} = \{E_{(0,0)}, E_{(1,0)}, E^{(1,0)}_{(0,0)}, E^{(1,0)}_{(1,0)}, E^{(0,1)}_{(0,0)}, E^{(0,1)}_{(1,0)}\}.$$

I.e., we consider function evaluation and x and y derivatives each at two points which (without loss of generality because of affine invariance) may be taken as the origin and a unit step in the x direction. At first thought we might hope to interpolate over this sixfold set with the attractive space $\mathcal{T}_2^2$ since it has dimension $\binom{2+2}{2} = 6$. However, if we apply the functionals $\mathcal{L}$ to the general element,

$$a_1x^2 + a_2y^2 + a_3xy + a_4x + a_5y + a_6,$$

of $\mathcal{T}_2^2$ and write out the resulting Gramian which multiplies the coefficient vector **a**, we obtain

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

Column two is all zeros so that the system is singular. We may feel perhaps that this arises from the asymmetry of node distribution with respect to the y axis, leading to elimination of the y^2 term from the representation so that the full basis of $\mathcal{T}_2^2$ is not available for this problem. The question then becomes; what alternate (and necessarily higher degree) basis function must be brought in? The following Newton-Hermite treatment shows that x^3 will fill the gap.

2.8 Coalescence Heuristics and the Two Point, Full First Order, Bivariate Hermite Problem. The art in applying the Newton-Hermite method lies in merging lines in order to specify derivatives. I use the image of a pair of parallel lines converging together at a point to create a derivative there in the direction perpendicular to the lines. In one dimension this is trivial. In two dimensions things are more complicated since multiple lines can bring in derivatives in two ways according to (12). Which ones should we choose? For our example, we may reason as follows. The two derivatives must be selected by the two different ways, else each point would require two each primary and cross lines which would bring in mixed derivatives. To economize on primary lines, these must pass through both points. This brings in the y derivatives. Then cross lines at each point must be introduced for function evaluation, and a pair of parallel cross lines at each point to get the x derivative. This results in the following collection, shown in Table 1, of lines, nodes, functionals and basis functions all indexed by $I^1_{(3,1)}$. We use the conventional subscript notation to represent partial differentiation of the function f being approximated.

Table 1. $I^1_{(3,1)}$ solution to the bivariate, two point, full first order problem

$\langle i,j\rangle$	H_i	p_i	H_{ij}	p_{ij}	z_{ij}	r,s,t	$L_{ij}(f)$	$u_{ij}(x,y)$
0 0	0−1 0	$-y$	1 0 0	x	(0,0)	0 0 0	$f(0,0)$	1
0 1			1 0−1	$x-1$	(1,0)	0 0 0	$f(1,0)$	x
0 2			1 0 0	x	(0,0)	0 1 0	$f_x(0,0)$	$x(x-1)$
0 3			1 0−1	$x-1$	(1,0)	0 1 0	$f_x(1,0)$	$x^2(x-1)$
1 0	0−1 0	$-y$	1 0 0	x	(0,0)	1 0 0	$f_y(0,0)$	$-y$
1 1			1 0−1	$x-1$	(1,0)	1 0 0	$f_y(1,0)$	$-yx$

The final column of the fourth row lets us know that of the four possible bivariate cubic monomials, x^3 is the one for us.

Cranking the handle on the recursion is the easiest part of the algorithm. It even produces a fairly self evident interpolant.

$$\begin{aligned}\pi f = q_{11}(x,y) = f(0,0)+&\\ &\big(f(0,0) - f(1,0)\big)x+\\ &\big(f(0,0) - f(1,0) - f_x(0,0)\big)x(x-1)+\\ &\big(-2f(0,0)+2f(1,0)+f_x(0,0)+f_x(1,0)\big)x^2(x-1)+\\ &f_y(0,0)y+\\ &\big(-f_y(0,0)+f_y(1,0)\big)xy.\end{aligned}$$

Sliding the last two rows four columns to the right shows how this expression mimics the sparse lower left triangular form of the Gramian.

I could continue this theme in pursuit of the Ženíšek elements, but it seems evident that some preliminary research in the Minsky direction would facilitate more productive numerical analysis.

3. Total Degree Newton-Hermite Systems

We noted that span $\{u_{ij}\}$ is included in a bivariate total degree space. Can conditions on the functional set be found so that the span of its Newton-Hermite basis is,

in fact, a total degree space? The answer is yes, and the condition which characterizes totality ensures that the total degree will be n, the number of primary lines in the functional specification. Gasca and Maeztu [1982] state that the characterization requires the secondary vector $\mathbf{m}$ to be of the form $\mathbf{m}^*$ where,

$$m_i^* = n - i, \quad i = 0, \ldots, n. \tag{15}$$

The sufficiency of (15) is obvious since in that case $\dim \operatorname{span}\{u_{ij}\} = \#I^n_{\mathbf{m}^*} = \binom{n+2}{2}$ and all basis functions are of total degree not exceding n. That it is necessary is a consequence of another multi-index

LEMMA: *If a Newton-Hermite index set has triangular size of* $\binom{n+2}{2}$ *and its elements are bounded (in the index 1-norm) by* n*, then the index set is* $I^n_* := I^n_{\mathbf{m}^*}$*. I.e.,*

$$\left\{ \begin{array}{c} \#I^{n'}_{\mathbf{m}} = \binom{n+2}{2} \\ \forall \iota \in I^{n'}_{\mathbf{m}} \Longrightarrow |\iota| \leq n \end{array} \right\} \Longrightarrow \left\{ \begin{array}{l} n' = n, \\ \mathbf{m} = \mathbf{m}^*. \end{array} \right\}$$

Proof: If $\forall \iota \ |\iota| \leq n$ then $i + m_i \leq n$, $i = 1, \ldots, n'$ and $m_i \geq 0 \Longrightarrow n' \leq n$. Suppose, contrary to the second consequence, that there is a $k \in \{0, \ldots, n'\}$ for which $k + m_k < n$. But we still have for all $i = 0, \ldots, n'$

$$m_i + 1 \leq n + 1 - i$$

so that

$$\sum_{i=0}^{n'}(m_i + 1) < \sum_{i=0}^{n'}(n+1) - \sum_{i=0}^{n'} i$$

$$(n+1)(n+2)/2 < (n'+1)(n+1) - n'(n'+1)/2 \tag{16}$$

$$n^2 - 2nn' + n'^2 < n' - n$$

$$0 \leq (n'-n)^2 < n' - n \leq 0$$

$$0 < 0 \qquad \gg\ll$$

This contradiction forces $\mathbf{m} = \mathbf{m}^*$. Then we have $m_i + 1 = n + 1 - i$, $i = 0, \ldots, n'$, so that '=' replaces '<' in (16). Then the conclusion is

$$0 \leq (n'-n)^2 = n' - n \leq 0$$

$$\Longrightarrow n' = n \qquad \blacksquare$$

Thus if we have a Newton-Hermite system $\{\langle H_i, \{H_{ij}\}\rangle \mid \langle i,j\rangle \in I^{n'}_{\mathrm{m}}\}$ specifying functionals by (12) and basis functions by (13), we know that $\deg u_{ij} = i+j$. Then if $\operatorname{span}\{u_{ij}\} = \mathcal{T}^2_n$, we must have $i+j \leq n$ and $\#I^{n'}_{\mathrm{m}} = \binom{n+2}{2}$. From the lemma we can deduce that $n' = n$ and $\mathrm{m} = \mathrm{m}^*$. The conclusion that Gasca and Maeztu [1982] draw is

THEOREM: *Arbitrary values are interpolated by $\mathcal{T}^2_n$ over a bivariate Newton-Hermite functional set if and only if the corresponding Newton-Hermite multi-index set is I^n_*.*

COROLLARY: *For a given point set $\{z_{ij}\}$ in the plane, if there is a set $\{H_i\}_{i=0}^n$ of $n+1$ lines in the plane such that $\{z_{ij}\} \subset \cup_{i=0}^n H_i$ and $\#(H_i \cap \{z_{ij}\}) = i+1$, then arbitrary values are interpolated by $\mathcal{T}^2_n$ over the Lagrange bivariate functional set $\{E_{z_{ij}} \mid \langle i,j\rangle \in I^n_*\}$.*

The set of lines will, of course, be the primary lines of the Newton-Hermite specifications of the functional set. Then the Gasca and Maeztu construction of §2.4 proves interpolation by the Newton-Hermite basis. By hypothesis of the corollary, the index set is I^n_*. Then the above theorem requires that the Newton-Hermite basis span $\mathcal{T}^2_n$.

Michelli [1986] and, following him, Cheney [1986] state the above corollary as a theorem which they prove somewhat indirectly by means of theorems concerning root sets of bivariate polynomials. Niether mentions the Gasca and Maeztu [1982] paper where the result (through considerably more proving effort, to be sure) is derived in greater Hermite generality.

APPENDIX I

Table 2. Fifth Total Degree Lagrange Interpolation in Four Dimensions

q-dim cell	$q+1$	ϕ	ψ	$\binom{m-1}{q}$	$\binom{n+1}{q+1}$	*#nodes*	$l_\phi(\mathbf{x}) = \prod_{k=1}^{5}(5x_{\phi_k} - \psi_k + 1)/\psi_k$
Hypersolid	5	*01234*	*11111*	1	1	1	$5^5\, x_0x_1x_2x_3x_4$
Solid	4	*ijkll*	*11112*	4	5	20	$5^4/2\, x_ix_jx_kx_l(5x_l - 1)$
		ijkkl	*11121*				$5^4/2\, x_ix_jx_k(5x_k - 1)x_l$
		ijjkl	*11211*				$5^4/2\, x_ix_j(5x_j - 1)x_kx_l$
		iijkl	*12111*				$5^4/2\, x_i(5x_i - 1)x_jx_kx_l$
Face	3	*ijkkk*	*11123*	6	10	60	$5^3/2\cdot 3\, x_ix_jx_k(5x_k - 1)(5x_k - 2)$
		ijjkk	*11212*				$5^3/2\cdot 2\, x_ix_j(5x_j - 1)x_k(5x_k - 1)$
		iijkk	*12112*				$5^3/2\cdot 2\, x_i(5x_i - 1)x_jx_k(5x_k - 1)$
		ijjjk	*11231*				$5^3/2\cdot 3\, x_ix_j(5x_j - 1)(5x_j - 2)x_k$
		iijjk	*12121*				$5^3/2\cdot 2\, x_i(5x_i - 1)x_j(5x_j - 1)x_k$
		iiijk	*12311*				$5^3/2\cdot 3\, x_i(5x_i - 1)(5x_i - 2)x_jx_k$
Edge	2	*ijjjj*	*11234*	4	10	40	$5^3/2\cdot 3\cdot 4\, x_ix_j(5x_j - 1)(5x_j - 2)(5x_j - 3)$
		iijjj	*12123*				$5^2/2\cdot 2\cdot 3\, x_i(5x_i - 1)x_j(5x_j - 1)(5x_j - 2)$
		iiijj	*12312*				$5^2/2\cdot 2\cdot 3\, x_i(5x_i - 1)(5x_i - 2)x_j(5x_j - 1)$
		iiiij	*12341*				$5^2/2\cdot 3\cdot 4\, x_i(5x_i - 1)(5x_i - 2)(5x_i - 3)x_j$
Vertex	1	*iiiii*	*12345*	1	5	5	$5/5!\, x_i(5x_i - 1)(5x_i - 2)(5x_i - 3)(5x_i - 4)$
				$2^{5-1} = 16$	31 $\|$ $2^5 - 1$	$126 = \binom{4+5}{4}$	

The nodes of U_5^4 are identified geometrically and grouped according to the dimension of the part of the simplex (column one) containing them. The second column is one more then the dimension of the subsimplex for which are listed the node identifiers ϕ and Lagrange polynomial. From the coordinate geometry, this is the number of positive elements in the $n+1$-tuple of barycentric node coordinates. Accordingly, it counts the number of distinct elements in each node identifier ϕ, and so the number of steps in ϕ. The ϕ in column three are to be taken as over all combinations of i, j, k, l such that $0 \le i < j < k < l \le 4$. The repeated elements of ϕ are counted by ψ in column four in an obvious way. The fifth column is the number of nodes which occur in the *interior* of each subsimplex (or on each main vertex). The index ψ is distinct for each of these but is the same for similarly positioned nodes on other like dimensioned subsimplices. The sixth column is the number of subsimplices of each dimension. This is also the number of node identifiers with the same ψ. Thus the product of column five times column six shown in column seven is the number of nodes of U_5^4 which occur in a subsimplex of given dimension. The last column shows the form of the corresponding node polynomial.

APPENDIX II

Infinite Interpolation: Countable Dimension

1. Introduction

Paul Halmos, speaking at the University of Alaska, Fairbanks in 1986, suggested that interesting distinctions of countably infinite logic can be found by extending results of linear algebra to infinite dimensional cases. Library research indicates that following this path through interpolation carries us into applications of several major theorems of twentieth century analysis including practical topology.

The functional analysis of Schauder bases and Hilbert space, as in Kreyszig [1973], provides a rigorous (although abstractly complex in its axiomatics) foundation for the study of infinite dimensional vector spaces. If we limit ourselves to countable dimension, the axiomatics are non-controversial and most likely consistent. The linear representation takes the form outlined in §I.3 and the fundamental theorem of §I.7 can be extended readily but not completely.

The non-singularity conditions b), e) carry over because of uniqueness of representation. Biorthonormality d) and spanning of the functional set f) are still sufficient for existence of interpolation by virtue of providing representation. Unique functional specification, however, does not carry over to unique interpolant selection unless $\mathcal{H}$ and $\mathcal{H}^*$ are isomorphic which is not always the case with infinite dimensional spaces. Condition c) must be replaced with non-zero requirements on determinants of infinite classes of submatrices of the infinite Gramian.

A notable result of this kind is due to Pólya, according to Davis [1975], who found a way to construct a solution to an infinite set of absolutely convergent equations in an infinite number of unknowns,

$$\sum_{k=1}^{\infty} a_{jk}x_k, \qquad j = 1, \ldots, \infty.$$

The determinant condition is that there exist a non-zero $n \times n$ determinant selectable from every infinite contiguous subblock of the first n rows of the Gramian. The further

condition is needed that the coefficients of an equation be asymptotically dominated by those of the following equation, i.e.,

$$\lim_{k\to\infty} a_{j-1,k}/a_{jk} = 0.$$

Under these circumstances there exists a solution for any infinite right hand side b_j, but not uniqueness. This loss is not surprising given the unending ability of infinite systems to soak up constructions without terminating. However, this extended multiplicity can also eliminate solvability. In another example from Davis [1975], we can see how the local Cauchy/Taylor power series representation of holomorphic functions,

$$f(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{n!} z^n,$$

might lead us to expect to match derivatives to given values at $z = 0$, i.e., to have infinite Hermite interpolation by analytic functions at a point. Unfortunately for such hopes, we cannot interpolate every infinite value set. For instance, $f^{(n)}(0) := (n!)^2$ grows too fast for the series to converge. Indeed, the conditions under which holomorphic functions will interpolate to fairly general value sets has been thoroughly investigated in modern complex analysis, and most of the issues have been settled.

2. Holomorphic Interpolation

It turns out that we can interpolate arbitrary values on sets satisfying a discreteness condition.

2.1 Uniqueness. This condition arises from a uniqueness theorem usually stated as an identity result, see, for instance, Rudin [1987a].

THEOREM: *A function holomorphic on an open connected set Ω is uniquely determined by its values on any subset of Ω which has a limit point in Ω.*

This does not say that a holomorphic function will match arbitrary values on such a set. Only that, at most, one will do so. Thus a point set with no limit points in the domain Ω will not be suitable for unique interpolation by holomorphic functions in general because it may be the root set of a holomorphic function on Ω. The identity theorem stated in terms of the root set becomes

THEOREM: *For any holomorphic f on an open connected set $\Omega \subset C$, the root set, $Z_f := f^{-1}(\{0\})$ is either Ω or has no limit points in Ω. In the latter of these two exclusive possibilities each of the at most countable number of roots, z_i, will be due to a factor $(z - z_i)^{m_i}$ in f and $(z - z_i)^{-m_i} f(z)$ will be non-zero at z_i.*

The result is essentially topological. We work in the relative topology of the domain Ω of f. The series representation gives the algebraic condition which shows that unless the function is identically zero in the neighborhood of a root, then it must be non-zero throughout a deleted neighborhood of that root. This means that all roots are either isolated or surrounded by a neighborhood of roots. Consider the derived set Z'_f of limit points of Z_f. By continuity of f it consists of roots. By definition it excludes isolated points and thus must, by the neighborhood condition just derived for roots, include a neighborhood of each of its points and so be an open set. But in any T_1 topological space the derived set is closed. Since the space Ω is connected we must have $Z'_f = \emptyset$ or(exclusively) $Z'_f = \Omega$. Since $Z'_f \subset Z_f$, the latter implies $Z_f = \Omega$, i.e., f is identically zero. Also $Z'_f = \emptyset$ implies that Z_f intersects any compact subset of Ω at most finitely. Then Z_f is at most countable since open sets of the plane are σ-compact.

2.2 Existence. We content ourselves with subuniqueness and consider *discrete* node sets (with no limit points). Davis [1975] gives the following theorems, respectively due to Guichard and derived from the above Pólya theorem.

THEOREM:

$$\Omega = C, \lim_{k\to\infty} |z_k| = \infty \quad \Longrightarrow \quad \forall \{w_k\}^\infty \subset C \; \exists f \in H(C) \quad f(z_k) = w_k, \quad \forall k.$$

THEOREM:

$$\Omega = \{|z| < r\}, \forall i \; |z_{i+1}| > |z_i|, \; \lim |z_i| = r \qquad \Longrightarrow$$
$$\forall \{w_{ij} \mid \langle i,j \rangle \in I^\infty_{\mathbf{m}}\} \; \exists f \in H(\Omega)$$
$$f^{(j)}(z_i) = w_{ij} \qquad j = 0, \ldots, m_i.$$

Rudin [1987a] draws the same conclusion, that the countable general contiguous Newton-Hermite interpolation problem is solvable by holomorphic functions, but with the somewhat more general hypotheses

$$\Omega \text{ open} \subset C, \; \{z_i\}_{i=1}^\infty \subset \Omega, \; \{z_i\}' \cap \Omega = \emptyset.$$

The no-limit-point condition is implied by the hypotheses of the first two theorems. Rudin assumes it explicitly. Pólya's theorem is proved constructively as are the others via intricate applications of the Mittag-Leffler and Weierstrasse factorization (or product) theorems which permit construction of holomorphic functions with given roots and poles on discrete point sets.

2.3 Uniqueness Recovered. We see that most practical node sets can be interpolated but that restraints on the node sets permit them to be zero sets of holomorphic functions so that interpolation solutions occur by the manifold. Imposing conditions on the node set and the interpolating function set such that uniqueness is attained has occupied a generation of very influential modern analysts. Over the years 1915 to 1932 a number of significant results were achieved by Blaschke, Carathéodory, Denjoy, Féjer, Gronwall, Nevanlinna, Pick, F. Riesz, Schur and J. L. Walsh. The subject of Blaschke Products — the basic form of the approximating function— forms the core material. The 1936 book by the last author is the definitive work (5^{th} ed. 1969) and contains an extensive bibliography. That the field is still very active can be seen from the monograph by P. Colwell [1985] reviewed by Rudin [1987b]. The results in this area lead us far from the linear setting so we will be very sketchy.

The condition on the function space is that it consist of bounded elements. Without loss of generality (Riemann Mapping Theorem) $\Omega = U := \{z \mid |z| < 1\}$. Let $H^p := \{f \in H(U) \mid \|f\|_p < \infty\}$. For $1 \leq p < \infty$ the interpolation problem can be solved. We quote Walsh [1969] for the case of H^2. Given a node set $\{z_k\} \subset U$ and an arbitrary value set, Walsh constructs a Newton-like representation over rational basis functions,

$$(1)\quad g(z) = \frac{a_0}{1-\overline{z}_1 z} + a_1 \frac{z - z_1}{(1-\overline{z}_1 z)(1-\overline{z}_2 z)} + a_2 \frac{(z-z_1)(z-z_2)}{(1-\overline{z}_1 z)(1-\overline{z}_2 z)(1-\overline{z}_3 z)} + \cdots,$$

and shows that this representation converges uniformly to an interpolant in H^2 provided that the series

$$\sum_{k=0}^{\infty} \frac{|a_k|^2}{1-|z_{k+1}|^2}$$

converges. Furthermore if $\prod^{\infty} |z_k|$ diverges (to zero), then the interpolant is unique. The product condition is equivalent, under the assumptions on z_k, to the convergence

of $\sum^{\infty} |1 - z_k|$. Note that this sum will diverge if $\{z_k\}$ has a limit point in U. If $\prod^{\infty} |z_k|$ converges then the representation (1) is still valid and defines an interpolant, but uniqueness is lost. The general form of the interpolant is then $g(z) + B(z)h(z)$, where h is general in H^2 and $B(z)$ is the Blaschke Product having roots at z_k,

$$B(z) := \prod_{k=1}^{\infty} \frac{\overline{z}_k}{|z_k|} \frac{z - z_k}{\overline{z}_k z - 1}.$$

It is also possible to have g and h from uniformly bounded function sets. Such is not, however, a linear space. The space H^{∞} has also been studied at length, but the results here are not nearly as straightforward. Existence comes constructively by means of a recursion over rational (Blaschke-like) functions. Uniqueness can be obtained by restricting interpolants to a uniformly bounded (without loss of generality, bounded by 1) family (not linear space) of functions. This condition is necessary but the value set must also satisfy certain further restrictions related to the recursion formulas. I refer the interested reader to Chapter 10 of Walsh [1969].

3. Univariate Countable Interpolation

Complex analysis is implicitly multivariate–wherin lies its complexity. Infinite interpolation is simpler in spaces of univariate functions.

3.1 Analytic Taylor-Hermite Interpolation. Define analytic functions on an open interval as those representable by convergent power series,

$$\mathcal{A}(r) := \{f \in R^{(-r,r)} \mid \exists \{a_k\} \in R^N \; \forall x \; f(x) = \sum_0^{\infty} a_k x^k\}$$

By Abel's theorem, if also $\sum a_k$ converges, then

$$\lim_{x \to r^-} f(x) = \sum^{\infty} a_k.$$

So define Abel functions,

$$\mathcal{A}'(r) := \{f \in \mathcal{A}(r) \mid \sum^{\infty} a_k < \infty\}.$$

Then every Abel function extends uniquely to a continuous function on $(0,r]$ with value $\sum a_k$ at r. Further consider functions differentiable at r,

$$\mathcal{LC}^\infty(r) := \{f \in R^{(-r,r)} \mid \forall n\ \exists \rho_n < r \quad f \in C^n(\rho_n, r), \quad \lim_{x \to r^-} f^{(n)}(x) = b_n\}.$$

Abel's theorem shows that $\mathcal{A}'(r)$ is a subspace of $\mathcal{LC}^\infty(r)$. It is proper since there are C^∞ functions which are not analytic. Finally define an Hermite functional set on $\mathcal{LC}^\infty(r)$ by

$$\mathcal{L}(r) := \{L_k \in (\mathcal{LC}^\infty(r))^* \mid L_k(f) := \lim_{x \to r^-} f^{(k)}(x)\}.$$

Then a theorem which Davis [1975] attributes to Borel becomes

THEOREM:

$$\forall f \in \mathcal{LC}^\infty(r) \quad \exists! g \in \mathcal{A}'(r) \quad \forall L \in \mathcal{L}(r) \qquad L(g) = L(f),$$

i.e., Abel functions interpolate locally C^∞ functions over differentiation of every order at the right endpoint. Here the functional set is countable.

3.2 Cardinal Harmonic Countable Interpolation. An example of greater utility is provided by the sampling theorem (as in H. J. Weaver [1983]) which states that band limited functions can be represented over a basis of sinc functions—moreover the representation is cardinal and orthogonal. Thus consider the space of summable signals with finite Fourier transforms which have convergent Fourier series,

$$\mathcal{F} := \{f \in C^R \mid f \in L^1(R),\ \hat{f} \in L^1(R),\ f \in \mathcal{NBV}\},$$

where $\hat{f} := \int_{-\infty}^{\infty} e^{i\omega t} f(t)dt$, is the Fourier transform, and $\mathcal{NBV}$ is normalized functions of bounded variation which therefore have Fourier series representation when replicated periodically. The approximating space will be signals of band limited spectrum,

$$\mathcal{H} := \{f \in \mathcal{F} \mid \omega \geq \omega_0 \implies \hat{f}(\omega) = 0\}.$$

The functional set is evaluation over a regular node set

$$\mathcal{L} := \{E_{k/\omega_0}\}_{k \in Z}.$$

the dual set biorthogonal to $\mathcal{L}$ gives the cardinal basis

$$l_k(t) = \operatorname{sinc}(\omega_0 t - k),$$

where $\operatorname{sinc} t = (\sin \pi t)/(\pi t)$ defined by continuity at $t = 0$. Then the sampling theorem amounts to

$$\mathcal{H} = \operatorname{span}\{l_k(t)\}_{k=-\infty}^{\infty}.$$

Now given any $f \in \mathcal{F}$, the function

$$g(t) = \sum_{k=-\infty}^{\infty} f(k/\omega_0) l_k(t)$$

is the unique element of $\mathcal{H}$ such that $g(k/\omega_0) = f(k/\omega_0)$, $\forall k \in Z$. In summary we may say that normalized summable functions of bounded variation with summable spectra are interpolated by band limited functions over a discrete regular sampling. It is not hard to show also that the sinc basis is orthogonal, $\int_{-\infty}^{\infty} l_j(t) l_k(t) dt = \delta_{jk}$.

This approach has been extended to the complex domain by E.T. Whittaker and later by J.M. Whittaker. We follow the review article by McNamee, Stenger and Whitney [1971], paraphrasing their representation result as follows. Without loss of generality we can normalize the grid spacing to unity so that the node set is the integers. Our subspace will be entire functions whose real sections are summable and whose imaginary sections are exponentially bounded,

$$\mathcal{H} := \{f \in H(C) \mid f|_R \in L^2(R),\ f(z) = \mathcal{O}(\exp \pi |\Im z|)\}.$$

Then elements of this space can be represented over a sinc basis,

$$\mathcal{H} = \operatorname{span}\{\operatorname{sinc}(z - k)\}_{k=-\infty}^{\infty},$$

with the cardinal representation,

$$f(z) = \sum_{k=-\infty}^{\infty} f(k) \operatorname{sinc}(z - k).$$

The essence of the series representation follows from an integral representation of elements of $\mathcal{H}$ given by the Paley-Wiener Theorem.

THEOREM: $\forall f \in \mathcal{H} \quad \exists F(\omega) \in L^2(-\pi, \pi)$

$$f(z) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i\omega z} F(\omega) d\omega.$$

4. Cardinal Spline Interpolation

No discussion of infinite interpolation could omit this elegant and very precisely laid out theory by I. J. Schoenberg [1973]. Briefly, we consider the approximating space of *cardinal splines* of degree n which are piecewise polynomials on each unit interval of the line,

$$\mathcal{S}_n := \{s(x) \mid s \in C^{n-1}(R),\ s|_{[k,k+1]} \in \mathcal{P}_n,\ k \in Z\}.$$

A convenient basis is provided by the *B-Splines*,

$$Q_n(x) := \frac{1}{(n-1)!} \sum_{k=0}^{n} (-1)^k \binom{n}{k} (x-k)_+^{n-1},$$

which give the representation for $s \in \mathcal{S}_{n-1}$,

$$s(x) = \sum_{k=-\infty}^{\infty} c_k Q_n(x-k).$$

The values of c_k can be explicitly computed in terms of s. The interpolation problem is then to find an $s \in \mathcal{S}_n$ which assumes arbitrary values over the integers. However, except for $n = 1$ the problem is underspecified and there is generally a manifold of solutions of dimension $n-1$. To achieve uniqueness we make the reasonable restriction of polynomial growth rates on both the value sets and the spline space. So define value sets

$$\mathcal{Y}_r := \{(y_k) \mid y_k = \mathcal{O}(|k|^r),\ k \in Z\},$$

and splines

$$\mathcal{S}_{n,r} := \{s \in \mathcal{S}_n \mid s(x) = \mathcal{O}(|x|^r)\}.$$

In our brevity we are presenting only half the picture. These splines will work for odd degree. A very symmetrical space S_n^* of splines with knots at half integer points solves the problem for even degree. Disregarding this split case structure, we can say that $S_{n,\tau}$ interpolates $\mathcal{Y}_\tau$ over $\{E_k\}_{k\in Z}$. To prove this, Schoenberg constructs a Lagrange basis of splines $l_n(x) \in S_n$ which are bounded and biorthonormal over the integers by virtue of $l_n(k) = \delta_{k,0}$. Then the unique spline solution of degree n to the interpolation problem is

$$s(x) = \sum_{k=-\infty}^{\infty} y_k l_n(x-k), \tag{2}$$

which converges absolutely and uniformly on any finite interval. These Lagrange basis functions are defined as follows. The splines,

$$M_{n+1}(x) := Q_{n+1}(x + \frac{n+1}{2}),$$

are such that the rational function in the complex variable t,

$$\sum_{k=-(m-1)}^{m-1} M_{2m}(k)t^k,$$

is different form zero throughout an annulus including the unit circle $|t| = 1$. Therefore we can expand its reciprocal in a Laurent series to get

$$1 \Big/ \sum_{k=-(m-1)}^{m-1} M_{2m}(k)t^k = \sum_{-\infty}^{\infty} d_k t^k.$$

Finally we define the cardinal basis functions

$$l_{2m-1}(x) := \sum_{k=-\infty}^{\infty} d_k M_{2m}(x-k),$$

which, amazingly, will do the job of representation in (2). Through further magic of this sort, Schoenberg is able to construct a cardinal representation of the solution by splines of a given degree to the m^{th}-order Lagrange-Hermite interpolation problem. This has the form

$$s(x) = \sum_{i=-\infty}^{\infty} \sum_{j=0}^{m} y_i^{(j)} l_j(x-i),$$

where $y_i^{(j)}$ are the values the j^{th} derivative is to assume at $x = i$.

REFERENCES

Garrett Birkhoff, "The algebra of multivariate interpolation," *Constructive Approaches to Mathematical Models,* Academic Press, Englewood Cliffs, New Jersey, pp. 345-363 [1978]

E. W. Cheney, "Multivariate approximation theory: Selected topics," *CBMS-NSF Regional Conference Series in Applied Mathematics,* SIAM Press [1986]

K. C. Chung and T. H. Yao, "On lattices admitting unique Lagrange interpolations," *SIAM J. Numer. Anal.,* v. 14, pp. 735-743 [1977]

P. Colwell, *Blaschke Products: Bounded Analytic Functions,* University of Michigan Press, Ann Arbor [1985]

S. A. Coons, "Surfaces for computer aided design of space forms," *Project MAC Report # MAC-TR-41,* Mass. Inst. of Tech., Cambridge, Massachusetts [1967]

P. J. Davis, *Interpolation and Approximation,* Dover, New York [1975]

C. de Boor, *A Practical Guide to Splines,* Springer-Verlag, New York [1978]

C. de Boor, "Topics in multivariate approximation theory," *MRC Technical Summary Report #2379,* University of Wisconson, Madison [1982]

A. R. Forrest, "Interactive interpolation and approximation by Beziér polynomials," *Computer Journal,* v. 15, pp. 71-79 [1972]

M. Gasca and J. I. Maeztu, "On Lagrange and Hermite interpolation in R^k," *Numer. Math.,* v. 39, pp. 1-14 [1982]

W. J. Gordon, "Blending-function methods of bivariate and multivariate interpolation and approximation," *SIAM J. Numer. Anal.,* v. 8, pp. 158-177 [1971]

A. S. Householder, *Principles of Numerical Analysis,* McGraw-Hill, New York [1953]

S. Karlin and W. J. Studden, *Tchebycheff Systems: with Applications in Analysis and Statistics,* Interscience, New York [1966]

S. Karlin, *Total Positivity,* Stanford University Press, Stanford, California [1968]

E. Kreyszig, *Introductory Functional Analysis with Applications,* John Wiley & Sons, New York [1973]

J. McNamee, F. Stenger and E. L. Whitney, "Whittaker's cardinal function in retrospect," *Math. Comp.,* v. 25, pp. 141-154 [1971]

C. A. Michelli, "Algebraic aspects of interpolation," *Proceedings of Symposia in Applied Mathematics,* American Mathematical Society, v. 36, pp. 81-102 [1986]

G. M. Nelson, "A method for interpolating scattered data based upon a minimum norm network," *Math. Comp.* v. 40, pp. 253-271 [1983]

E. D. Nering, *Linear Algebra and Matrix Theory*, John Wiley & Sons, New York [1963]

R. A. Nicolaides, "On a class of finite elements generated by Lagrange interpolation," *SIAM J. Numer. Anal.*, v. 9, pp. 435-445 [1972]

C. Olmsted, "Two formulas for the general multivariate polynomial which interpolates a regular grid on a simplex," *Math. Comp.*, v. 47, pp. 275-284 [1986]

W. Rudin, *Real and Complex Analysis*, 3rd Ed., McGraw-Hill, New York [1987a]

W. Rudin, Review of P. Colwell [1985], *The American Mathematical Monthly*, v. 94, pp. 311-314 [1987b]

I. J. Schoenberg, "Cardinal spline interpolation," *CBMS-NSF Regional Conference Series in Applied Mathematics*, SIAM Press [1973]

I. J. Schoenberg, *Mathematical Time Exposures*, Mathematical Association of America [1982]

I. J. Schoenberg and A. Whitney, "On Pólya Frequency Functions III: The positivity of translation determinants with application to the interpolation problem by spline curves," *TAMS*, v. 74, pp. 146-259 [1953]

H. C. Thacher, Jr., "Generalization of concepts related to linear dependence," *J. Soc. Indust. Appl. Math.*, v. 6, pp. 288-299 [1958]

H. C. Thacher, Jr., "Derivation of interpolation formulas in several independent variables," *Ann. New York Acad. Sci.*, v. 86, pp. 758-775 [1960a]

H. C. Thacher, Jr. and W. E. Milne, "Interpolation in several variables," *J. Soc. Indust. Appl. Math.*, v. 8, pp. 33-42 [1960b]

J. L. Walsh, *Interpolation and Approximation by Rational Functions in the Complex Domain*, 5th Ed., American Mathematical Society, Providence, Rhode Island [1969]

H. J. Weaver, *Applications of Discrete and Continuous Fourier Analysis*, John Wiley & Sons, New York [1983]

A. Ženíšek, "Interpolation polynomials on the triangle," *Numer. Math.*, v. 15, pp. 283-296 [1970]

A. Ženíšek, "Hermite interpolation on simplices in the finite element method," *Proc. of EQUADIFF III*, Eds. Rab and Vosmansky, J. E. Purkyně U., Brno, pp. 271-277 [1973]